

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 03:58:17 ; Search time 11025 Seconds
(without alignments)
17546.947 Million cell updates/sec

Title: US-09-922-549B-67
Perfect score: 11945
Sequence: 1 tgcgcctctcttgatattc.....tgcgcctgcgtccttaggg 11945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estcov:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 630.4 | 5.3 | 699 | 9 | AL584990 AL584990 |
| 2 | 262.6 | 2.2 | 673 | 9 | AJ453074 AJ453074 |
| 3 | 238.6 | 2.0 | 634 | 13 | BI064896 BI064896 |
| 4 | 222.2 | 1.9 | 446 | 9 | AL587961 AL587961 |
| 5 | 220.8 | 1.8 | 582 | 12 | BG625446 BG625446 |
| 6 | 210.6 | 1.8 | 650 | 9 | AJ453847 AJ453847 |

| | | | | | |
|----|-------|-----|-----|----|---------------------|
| 7 | 208.6 | 1.7 | 689 | 9 | AJ392780 AJ392780 |
| 8 | 201 | 1.7 | 489 | 10 | AM240181 AM240181 |
| 9 | 197.4 | 1.7 | 700 | 9 | AJ445833 AJ445833 |
| 10 | 197.4 | 1.7 | 826 | 9 | AJ456416 AJ456416 |
| 11 | 193.2 | 1.6 | 914 | 17 | GG246112 GG246112 |
| 12 | 191 | 1.6 | 437 | 12 | BG624980 BG624980 |
| 13 | 186.4 | 1.6 | 608 | 17 | GG0000392 GG0000392 |
| 14 | 183.2 | 1.5 | 604 | 14 | B0037358 B0037358 |
| 15 | 183 | 1.5 | 581 | 13 | BI388759 BI388759 |
| 16 | 178.8 | 1.5 | 582 | 9 | AL588808 AL588808 |
| 17 | 175 | 1.5 | 624 | 13 | BM426214 BM426214 |
| 18 | 164.6 | 1.4 | 644 | 9 | AL585302 AL585302 |
| 19 | 164.2 | 1.4 | 681 | 9 | AJ393361 AJ393361 |
| 20 | 160 | 1.3 | 564 | 13 | BI066024 BI066024 |
| 21 | 158.4 | 1.3 | 502 | 13 | BM490080 BM490080 |
| 22 | 155 | 1.3 | 626 | 9 | AJ455303 AJ455303 |
| 23 | 154.8 | 1.3 | 644 | 13 | BI392307 BI392307 |
| 24 | 154.4 | 1.3 | 741 | 9 | AL584343 AL584343 |
| 25 | 154.2 | 1.3 | 643 | 13 | BM440491 BM440491 |
| 26 | 152.4 | 1.3 | 637 | 13 | BM440335 BM440335 |
| 27 | 151.4 | 1.3 | 900 | 14 | B0720191 B0720191 |
| 28 | 149 | 1.2 | 644 | 13 | BI390882 BI390882 |
| 29 | 148.4 | 1.2 | 647 | 13 | BM440702 BM440702 |
| 30 | 148.2 | 1.2 | 796 | 9 | AJ444515 AJ444515 |
| 31 | 146.4 | 1.2 | 604 | 17 | GG4140117 GG4140117 |
| 32 | 144.4 | 1.2 | 634 | 13 | BM486806 BM486806 |
| 33 | 144.2 | 1.2 | 562 | 13 | BM426795 BM426795 |
| 34 | 143.4 | 1.2 | 483 | 13 | BM489773 BM489773 |
| 35 | 143 | 1.2 | 645 | 13 | BM488238 BM488238 |
| 36 | 142.8 | 1.2 | 455 | 17 | BH405358 BH405358 |
| 37 | 140.4 | 1.2 | 664 | 13 | BI393232 BI393232 |
| 38 | 139.6 | 1.2 | 754 | 9 | AJ454778 AJ454778 |
| 39 | 139 | 1.2 | 620 | 14 | B0038729 B0038729 |
| 40 | 138.8 | 1.2 | 600 | 17 | GG0000010 GG0000010 |
| 41 | 138.4 | 1.2 | 804 | 9 | AJ445375 AJ445375 |
| 42 | 137.2 | 1.1 | 638 | 13 | BM440658 BM440658 |
| 43 | 135.6 | 1.1 | 676 | 13 | BM426719 BM426719 |
| 44 | 134.6 | 1.1 | 644 | 9 | AJ981722 AJ981722 |
| 45 | 134.6 | 1.1 | 676 | 10 | AM983878 AM983878 |

ALIGNMENTS

| | | | | | |
|------------|---|--|------|--------|-----------------|
| RESULT 1 | AL584990 | 699 bp | mRNA | linear | EST 28-FEB-2001 |
| LOCUS | AL584990 | Stratagene Chick Embryo Lambda cDNA Library | | | |
| DEFINITION | AL584990 | Gallus gallus cDNA clone ROS017A10, mRNA sequence. | | | |
| ACCESSION | AL584990.1 | GI:1163723 | | | |
| VERSION | EST. | | | | |
| KEYWORDS | ORGANISM | chicken | | | |
| SOURCE | Gallus gallus | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauia; Aves; Neognathae; Galliformes; Phasianidae; | | | | |
| REFERENCE | 1 (bases 1 to 699) | | | | |
| AUTHORS | Murray, F. | | | | |
| TITLE | Stratagene Chick Embryo Lambda cDNA Library | | | | |
| JOURNAL | Unpublished (2001) | | | | |
| COMMENT | Contact: Frazer Murray | | | | |
| | Dept. Genomics and Bioinformatics | | | | |
| | Roslin Institute | | | | |
| | Roslin, Midlothian, EH25 9PS, UK | | | | |
| | Tel: +44 (0)131 527 4200 | | | | |
| | Fax: +44 (0)131 440 0434 | | | | |
| | Email: frazer.murray@bbsrc.ac.uk | | | | |
| | Seq primer: 73. | | | | |
| FEATURES | location/Qualifiers | | | | |
| source | 1..699 | | | | |
| | /organism="Gallus gallus" | | | | |
| | /db_xref="taxon:9031" | | | | |

| KEYWORDS | EST. |
|-----------------------|--|
| SOURCE | Chicken. |
| ORGANISM | Gallus gallus |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus. 1 (bases 1 to 634) |
| AUTHORS | Cogburn, L.A., Morgan, R.W. and Burnside, J. |
| TITLE | Chicken ESTs from fat |
| JOURNAL | Unpublished (2001) |
| COMMENT | Contact: Larry A. Cogburn University of Delaware Townsend Hall, Newark, DE 19717, USA Tel: 302-831-1335 Fax: 302-831-2822 Email: cogburn@udel.edu, www.chickest.udel.edu. |
| FEATURES | Location/Qualifiers |
| source | 1..634 /organism="Gallus gallus" /db_xref="taxon:9031" /clone_pfln:"PK002.n24" /clone_lib="normalized chicken fat cDNA library" /sex="Male and Female" /tissue_type="fat" /lab_host="E.coli EMDH10B" /note="Vector: pSPORI1" |
| BASE COUNT | 139 a 186 c 122 g 178 t 9 others |
| ORIGIN | |
| Query Match | 2.0%; Score 238.6; DB 13; Length 634; |
| Best Local Similarity | 67.7%; Pred. No. 1,1e-30; |
| Matches | 388; Conservative 0; Mismatches 180; Indels 5; Gaps 4; |
| OY | 10613 GGATCATGAAGATCCACACCCTGGCAGAGGCCACCAACTTCAGATCGTAGC 10672 |
| Db | 1 GGATCATGAAGTGCACACCCCCTTGCTGACAGAAAGGTGTCACTGCAATCAAGCAC 60 |
| OY | 10673 TAGACCAAGGAGCCAGGCGTCCATCCAAACCGGCATGANACCTCCAGGATGAGCA 10732 |
| Db | 61 TGGGGAGAGTTCCCATGGTCCCATGCACAGCCGTGGATGACCTTGGGGATGGCCA 120 |
| OY | 10733 TCACACAACCTCTCTGTGGAGAGCCTGTGCCAGCAGCTCA-CACACTCTGTGTGAAGAATT 10791 |
| Db | 121 TCACAGAGCCTCTGGACAACCTGTGCCAGAGGTGTACACCACCTCTCAGTGGAAAACT 180 |
| OY | 10792 TTCCCTGACATCCATCTRAAGCTTCCCTCTTGAGGTAGATFCCACTCCCCCTGTGCT 10851 |
| Db | 181 TTGCTGACATCTAATTAATAATTTCCCTCTTGATTTAAACATCATCTCTCTGTGCG 240 |
| OY | 10852 ATCACTGTCT-ACCTTGTGAAGAAGTGAATGTCCTCTCTTTTGAAGGTGCATGAGG 10910 |
| Db | 241 ATTGCTGTCTGTTTTAAAAAAAAAAAAAACTTCTCTTCTGTTTAATTCCTCTTAAAG 300 |
| OY | 10911 TCTCCTTCAGAGCTCTTCTCTTCTTGACAGGATGAACAAGCCAGCTCCCTCAGCCGTCT 10970 |
| Db | 301 TATTGTAGGCAATTACAAAGTCTCCCAACAGTACAAAGCCAGCTTCCTCAGCCGTCT 360 |
| OY | 10971 TTATAGAGAGAGTGTCTCCAGCCCTTGATCATCTTTGTGGCCCTCTCTGGACCCGCTCC 11030 |
| Db | 361 TCATTTGAGAGATGCTCCACCTTCTCATCATCTTTGTGGCCCTCTC-AGACCCCTCTCC 419 |
| OY | 11031 AAGACCTCCACATCTTCTCTGTAGTGGGGGCCAGGCGCTGAATGACAGATCTCAGATGG 11090 |
| Db | 420 AAAAGCTCAGTATCT--CTGTGTAGTGTGGCCCAAGACTGATGATGACACCAAGATGG 477 |
| OY | 11091 GGCCTCAAAGAAGCAGAGTAAAGAGGAGCAATCACTTCTCACCCTGTGGCCAGCCCT 11150 |
| Db | 478 GGCCTCATATATGTGTAGAGAGAGAGAGATGCTCACCTCCCTCATCAGAGCTGGCCACTCT 537 |
| OY | 11151 CTTTGTATGAGAGCCCTGATTAACAATCGGCTTTC 11183 |
| Db | 538 TTTTGTATGTNNNCCNNNGTACTGTGTCTGTCG 570 |

[illegible]

| | | | |
|-----------------------|-------|---|-------|
| OY | 1187 | TTCCTACCCCTGCTGGCCAGGCCCTTGTGCAGTGAGCGCCTGGATTCACACTGCTTTCTGA | 11186 |
| Db | 357 | TMCCTCTCCCTGCTGCTGGCCACCCTTTTAATGACGCCAGCCAGACAGTGTGGCTTCCAG | 416 |
| OY | 1187 | GCTGCAA | 11193 |
| Db | 417 | GCTGCAA | 423 |
| <hr/> | | | |
| RESULT 5 | | | |
| BG625446/c | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| FEATURES | | | |
| source | | | |
| BASE COUNT | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches | | | |
| OY | 10570 | AATCATAGAAATCAATGAAATGGCTGCTGGTGTGGAGGAGGCCACCAAGATCATGAGATCCAA | 10629 |
| Db | 337 | AATCATAGAATCACAGAAATGGCTGCTGGTGTGGAGGAGGCCACCAAGATCATGAGATCCAA | 278 |
| OY | 10630 | CACCCCCCAGAGGAGGAGGAGGCCACCAACCTCCAGATCTGCTAGAC-CAGGAGGCCA | 10688 |
| Db | 277 | CTCCCCCAGCAGGAGGAGGCCACCAACCTCCCTCCATTAACTAGACTCAGGCTGCCA | 218 |
| OY | 10689 | GGGCTCATCCAACCTGGCCATGAAACACCTCCAGGAGATGAGACATCCACAACCTCTTG | 10748 |
| Db | 217 | GGCCCCCATCCAAGCGCGCTTGAACACCTTCAGGAGATGGGCGATCCACAACCTCTTG | 158 |
| OY | 10749 | GCAGCCTGTGCAGACCTCACACACCTCTCTGTGAAGACTTTTCCCTGCATCCCAATC | 10808 |
| Db | 157 | GCAGCCCTTCCAGACCTCACACCTCTCTGTGAAGACTTTTCCCTGTGACAAACAAC | 98 |
| OY | 10809 | TAAAGCTTCCCTCTGAGTGTAGATTCACATCCCCCTGTGCTATCAGTCTACTACTTG | 10868 |
| Db | 97 | TAAATCTTCCCTCCCTCAACTTAAACACACTTCCCCTGTCTCGCATTAATTCACGCTTG | 38 |
| OY | 10869 | TAAAAAGTTGATTC | 10884 |

| DB | 37 | TAAAGCTTGACTCCC | 22 |
|---------------------------|---|--------------------|-----------------|
| RESULT 6 | | | |
| LOCUS | AJ453847 | 650 bp | EST 22-APR-2002 |
| DEFINITION | AJ453847 riken1 Gallus gallus | CDNA clone 35m1orl | mRNA sequence. |
| ACCESSION | AJ453847 | | |
| VERSION | AJ453847.1 | GI:20263943 | |
| KEYWORDS | EST. | | |
| SOURCE | chicken. | | |
| ORGANISM | Gallus gallus | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | Archosauria; Aves; Neognathae; Galliformes; Phasianidae; | | |
| REFERENCE | | | |
| AUTHORS | Phasianine; Gallus. | | |
| TITLE | 1 (bases 1 to 650) | | |
| JOURNAL | Buerstedde, J.M. | | |
| COMMENT | Gallus gallus bursal lymphocyte EST | | |
| | unpublished (2002) | | |
| | Contact: Buerstedde JM | | |
| | Cellular Immunology | | |
| | Heinrich-Pette-Institute | | |
| | Marinistr. 52, 20251 Hamburg, Germany | | |
| | Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html . | | |
| FEATURES | | | |
| source | 1..650 | | |
| | /organism="Gallus gallus" | | |
| | /db_xref="taxon:9031" | | |
| | /clone="35m1orl" | | |
| | /clone_1lb="riken1" | | |
| | /cell_type="bursal lymphocyte" | | |
| | /dev_stage="2-3 weeks old" | | |
| | /note="CB inbred strain" | | |
| BASE COUNT | 129 a 208 c 166 g 147 t | | |
| ORIGIN | | | |
| Query Match | 1.8%; Score 210.6; DB 9; Length 650; | | |
| Best Local Similarity | 76.4%; Pred. No. 6.2e-26; | | |
| Matches 334; Conservative | 0; Mismatches 74; Indels 29; Gaps 5; | | |
| OY 10577 | GAATCATAGATGCGCTGGTTGGAAGGACCCCAAGATCATGAATCCACACCCC | 10636 | |
| Db 220 | GAATCATAGAAATGGCTGGTTGGAAGGACCTCAAGATCATGAATCCACACCCCCT | 279 | |
| OY 10637 | GCCACAGGCGGGGCCACCACTCCAGATCTGTGTCATGACAGGACGCCAGGGCTCCA | 10696 | |
| Db 280 | GC--TGGCAGGGGCCACCACTCCCAT--TTACTAGACCAAGTTGCCAGGGCCCCA | 334 | |
| OY 10697 | TCCAACTGGCCCATGAACACTCCAGGGATGGAGCATCCACACCTCTGTGGGACGCTG | 10756 | |
| Db 335 | TCCAACTGGCCCTTGAACACTCTAGGGATGGAGCATCCACACCTCTGTGGGACGCG | 394 | |
| OY 10757 | TGCCAGACCTCACCACCTCTCTGTGAAGAATTTCCCTGACATCCAAATCTAAGCTT | 10816 | |
| Db 395 | TTCCAGACCTCACCACCTCTCTGTGAAGAAG--TTCCCTAATCCAAATCTAATCTT | 453 | |
| OY 10817 | CCCTCTTGAAGTATGATCACTCCCTTGGCTATCATGTCATCTTGTAAAGT | 10876 | |
| Db 454 | CCCTCTTGAAGTATGATCACTTGTCCCTTGTCTGCTGTATCTACACCTTTCAAGAGT | 513 | |
| OY 10877 | TGATTCCTCTCTCTT-----TGAAGTTGCATGAGGTCTTCT | 10916 | |
| Db 514 | TTACTTCCTCTCTCTTTATAGGCTCATTTACAGGTACTGAAGGCTGAATGAGGTACCG | 573 | |
| OY 10917 | TGCACCTTCTCTTCTTCTGACAGATGAACAAGCCACCTCCCTCAGCTTCTTATAGT | 10976 | |
| Db 574 | GCGACAC--CTTCTCTCTCTCCAGGCTGAATAAGCCACCTCCCTCAGCTTCTTATAGT | 630 | |
| OY 10977 | GAGAGTGTCTCCAGCC 10993 | | |
| Db 631 | ACTAGGTATGTGTTC 647 | | |

[illegible]

| | | | |
|---------------------------|---|--|--------------------|
| OY | 7048 | TGATGACGACTTCCAGATGGGGCCTTA | 7075 |
| | | | |
| Db | 659 | TGATGACGACTTCCAGATGGAGNTTCA | 686 |
| RESULT 8 | | | |
| AM240181/c | | | |
| LOCUS | AM240181 | 489 bp | mrna |
| DEFINITION | pt1c.pk002.n15 chicken T cell cDNA library | linear | EST 27-APR-2000 |
| ACCESSION | AM240181 | | |
| VERSION | AM240181.1 | GI:6579921 | |
| KEYWORDS | EST. | | |
| SOURCE | chicken. | | |
| ORGANISM | Gallus gallus | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | Archosauria; Aves; Neognathae; Galliformes; Phasianidae; | | |
| | Phasianinae; Gallus. | | |
| | 1 (bases 1 to 489) | | |
| REFERENCE | | | |
| AUTHORS | Morgan, R. | | |
| TITLE | Chicken T cell ESTs | | |
| JOURNAL | Unpublished (1999) | | |
| COMMENT | Contact: Robin Morgan | | |
| | University of Delaware | | |
| | Townsend Hall, Newark, DE 19717, USA | | |
| | Tel: 302-831-1341 | | |
| | Fax: 302-831-2822 | | |
| | Email: morgan@udel.edu | | |
| | Clones can be ordered online at http://www.chickest.udel.edu . | | |
| FEATURES | | | |
| SOURCE | location/Qualifiers | | |
| | 1..489 | | |
| | /organism="Gallus gallus" | | |
| | /db_xref="taxon:9031" | | |
| | /clone="pt1c.pk002.n15" | | |
| | /clone_lib="chicken T cell cDNA library" | | |
| | /sex="Male and female" | | |
| | /cell_type="splenic T cell" | | |
| | /lab_host="E.coli DH10B" | | |
| | /note="Vector: pB42AD" | | |
| BASE COUNT | 134 a 87 c 135 g 106 t | 27 others | |
| ORIGIN | | | |
| Query Match | 1.7% | Score 201; | DB 10: Length 489; |
| Best Local Similarity | 67.7% | Pred. No. 3.1e-24; | |
| Matches 319; Conservative | 0; Mismatches 127; Indels 25; Gaps 3; | | |
| OY | 10660 | CCAGATCTGACTGACGAGCAGCCGAGGCGTCATCAACCTGGCCAT--GAACACC | 10717 |
| | | | |
| Db | 469 | CCCAATTAAATATTGACACAGAGTNCGCCAGCCCATCACTGACCTTGAACATTT | 410 |
| OY | 10718 | TCCAGGATGAGCATCCACAACCTCTCTGGGACGCTGTGCGACACCTACACCCCTC | 10777 |
| | | | |
| Db | 409 | TCCAGGGACAGGNCATCCACAACCTTTTGGGACAGCTTTCACACCTTACCACANTTT | 350 |
| OY | 10778 | TCTGTGAGAACCTTTCCTCGATCCCATCAATCAACCTTCCTCTGTGAGTTAGATCCA | 10837 |
| | | | |
| Db | 349 | TNATGTGAAGACTTCCTNMAACATCCAACTTAATNTTCCCTCCCTCAACTTAACCA | 290 |
| OY | 10838 | CTCCCCCTGTGCTATCACTGCTACTCTGTGTAAAGTTGATTCCTCCCTTTT----- | 10892 |
| | | | |
| Db | 289 | TTTCCCTTGTCNTGTTGTATCTACCCCTTNCMAAGAGGTGACCTCCCTCMTGTTATAG | 230 |
| OY | 10893 | -----TGAAGTTTGCATGAGGTCTCTTGACAGCTTCTTCTCTTCCG | 10937 |
| | | | |
| Db | 229 | GCTCCCTTAGGTANTGAAGGTTTGCATTTAGTGTACCCGGCACTTTT---TCTTCC | 173 |
| OY | 10938 | AGGATGAACAAGCCAGAGTCCCTCAGCCTGTCTTATAGAGAGGTGCTCCAGCCCTGTG | 10997 |
| | | | |
| Db | 172 | GGTTTGAACAAGCCAGAGTCCCTCAGCMTGTTTGTATGAGGAGGTGCTCCAGCATTTTG | 113 |
| OY | 10998 | ATCATCTTTGTGGCCCTCTCTTGAGCCGCGTCCCAAGAGCTCCACATCTTTCTGTAGTGG | 11057 |
| | | | |

| | | | |
|---|--|---|-------|
| Df | 112 | ATCATCTTTGAGCCCTCCTNTNGATTCCTGCCACACAGTCCCTGTATTNTGTAAATGG | 53 |
| OY | 11058 | GGGCCCCGAGCGCTGAATGCAGTACTCCAGATGGGGCCCTCAAAAAGCAGAG | 11108 |
| Df | 52 | GGGGCTCCAGAGCTTGACACAGTACTCCAGAAAAGNCCCACANGAGCAGAG | 2 |
| RESULT | 9 | | |
| LOCUS | AJ445833 | | |
| DEFINITION | AJ445833 rikenl Gallus gallus cdna clone l1j14r1, mRNA sequence. | | |
| ACCESSION | AJ445833 | | |
| VERSION | AJ445833.1 GI:20213054 | | |
| KEYWORDS | EST. | | |
| SOURCE | chicken. | | |
| ORGANISM | Gallus gallus | | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| Archosauria; Aves; Neognathae; Galliformes; Phasianidae; | | | |
| Phasianinae; Gallus. | | | |
| 1 (bases 1 to 700) | | | |
| Buerstedde,U.M. | | | |
| Gallus gallus bursal lymphocyte EST | | | |
| Unpublished (2002) | | | |
| Contact: Buerstedde JM | | | |
| Cellular Immunology | | | |
| Heinrich-Pette-Institute | | | |
| Martinstr. 52, 20251 Hamburg, Germany | | | |
| Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html. | | | |
| Location/Qualifiers | | | |
| 1..700 | | | |
| /organism="Gallus gallus" | | | |
| /db_xref="taxon:9031" | | | |
| /clone="l1j14r1" | | | |
| /clone_lib="rikenl" | | | |
| /clone_type="bursal lymphocyte" | | | |
| /dev_stage="2-3 weeks old" | | | |
| /note="CB inbred strain" | | | |
| BASE COUNT | 154 a 205 c 130 g 210 t | 1 others | |
| ORIGIN | | | |
| Query Match | 1.7%; Score 197.4; DB 9; Length 700; | | |
| Best Local Similarity | 68.5%; Pred. No. 1.1e-23; | | |
| Matches 316; Conservative 0; Mismatches 141; Indels 4; Gaps 3; | | | |
| OY | 10724 | GATGAGCATCCACACACCTCTCTGGGAGCCCTGTGCCACAGACTCA-CGACCCTCTGT | 10782 |
| Df | 1 | GATGGGCATCCACAGCCCTCTGGACACCTCTTCCCGAGGGCTCACCCACCCCTCTCAGT | 60 |
| OY | 10783 | GAGAAGCTTTCOCGACATCCATCTAAGCTTCCTCCCTGAGGTGAGATCCACCTCC | 10842 |
| Df | 61 | GGAAAACTTGGCCGACATCTAATCTAATTTTCCCTCTTAAGTTTAAACVATTCCT | 120 |
| OY | 10843 | CCTTGTGCTATCACTGCTACTCTTGTAAAGTAGTTCCTCTTGTGGAGAGTTG | 10902 |
| Df | 121 | CCTTGTGCGATTCGTCTGTTTAAAAAAAAAAAACTTCTTCTTGTATATATTC | 180 |
| OY | 10903 | CAATGAGCTCTCCCTTGCAGCCTTCTCTCTCTCTGAGATGAACAAGCCACCTCCCTCA | 10962 |
| Df | 181 | CCTTAAATATGTGAGGCACTTACAAAGTCTCCCAAGTCAACAAGCCACCTCTCTCA | 240 |
| OY | 10963 | GCCTGTCTTATAGAGAGGAGCTCCAGCCCTCTGATATCTTGTGGCCCTCTCTGGA | 11022 |
| Df | 241 | GCCTGTCTTCAATGAGAGATGCTCCACCTCTCATCATCTGTGTGGCCCTCTC -AGA | 299 |
| OY | 11023 | CCCGTTCAGAGGCTCCACATCTTCCGTAAGTGGGGGCCCGAGCCGTAATGACGTACT | 11082 |
| Df | 300 | CCCTCTCCAAGGCTCAGTATCTT--CTGTGATGTGGGCCCGAGCCTGAGTGCATACA | 357 |
| OY | 11083 | CCAGATGGGGCTCAAAAGAGCAGAGTAAGAGGAGACATCACTTCCCTCAGCCCTGG | 11142 |
| Df | 358 | CCAGATGGGGCTCAATAATGTGAGAGTGAAGAGGAGTACCTACCTCCCTCATCAGCTGG | 417 |
| OY | 11143 | CCAGCCCTCTCTGATVGAGCCCTGGATACCACTGGCTTTC | 11183 |

| DB | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM |
|----|-------|---|---|---|----------|-----------------|----------|
| Db | 418 | CCACCTCTTTTGTATGATGACCCAGGAGTACTGTTCTCTGC | 458 | | | | |
| | | RESULT 10 | | | | | |
| | | AJ456416 | 826 bp | MRNA | linear | EST 22-APR-2002 | |
| | | AJ456416 | | | | | |
| | | DEFINITION | AJ456416 | riken1 Gallus gallus gallus cDNA clone 8118r1, mRNA sequence. | | | |
| | | ACCESSION | AJ456416 | | | | |
| | | VERSION | AJ456416.1 | GI:20266512 | | | |
| | | KEYWORDS | EST. | | | | |
| | | SOURCE | chicken. | | | | |
| | | ORGANISM | Gallus gallus | | | | |
| | | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | | | Archosauria; Aves; Neognathae; Galliformes; Phasianidae; | | | | |
| | | | Phasianinae; Gallus. | | | | |
| | | | 1 (bases 1 to 826) | | | | |
| | | REFERENCE | Buerstedde, J.M. | | | | |
| | | AUTHORS | Gallus gallus bursal Lymphocyte EST | | | | |
| | | TITLE | Unpublished (2002) | | | | |
| | | JOURNAL | Contact: Buerstedde JM | | | | |
| | | COMMENT | Cellular Immunology | | | | |
| | | | Heinrich-Pette-Institute | | | | |
| | | | Martinistr. 52, 20251 Hamburg, Germany | | | | |
| | | | Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html . | | | | |
| | | FEATURES | Location/Qualifiers | | | | |
| | | source | 1. 826 | | | | |
| | | | /organism="Gallus gallus" | | | | |
| | | | /db_xref="taxon:9031" | | | | |
| | | | /clone="8118r1" | | | | |
| | | | /clone_id="riken1" | | | | |
| | | | /cell_type="bursal Lymphocyte" | | | | |
| | | | /dev_stage="2-3 weeks old" | | | | |
| | | | /note="CB inbred strain" | | | | |
| | | BASE COUNT | 191 a 234 c 160 g 240 t | | | | |
| | | ORIGIN | 1 others | | | | |
| | | Query Match | 1.7%; Score 197.4; DB 9; Length 826; | | | | |
| | | Best Local Similarity | 68.5%; Pred. No. 9.8e-24; | | | | |
| | | Matches 316; Conservative | 0; Mismatches 141; Indels 4; Gaps 3; | | | | |
| | | OY 10724 | GATGAGCATCCACACACCTCTCTGTGGGACCGCTGTGCCACGACCTCA-CCACCTCTCTGT | 10782 | | | |
| | | Db | 1 GATGGGCCATCCACACCGCTCTGGACACCTGTCCCGAGGATCTCACCCACCTCTCAGT | 60 | | | |
| | | OY 10783 | GAGAGACTTTTCCCGCATCCCAATCAAGCTCCCTCCCTGAGGTAGATCCACCTCC | 10842 | | | |
| | | Db | 61 GGAATAACTTTTGCTCGACATCTAATCTAAATTTTCCCTCTTGTAGTTAAACTATTTCT | 120 | | | |
| | | OY 10843 | CCCTGTGCTATCACTGTCTACTCTTGTAAAGAGTTGATTCCTCTTTTGGAGGTTG | 10902 | | | |
| | | Db | 121 CCTGTCCGATTTGCTGTCTGTTTAAAAAATACTTTCTTGTATTAATATTC | 180 | | | |
| | | OY 10903 | CAATGAGGTCTCTTGACGCTTCTTCTCTTCTGACAGATGAACAAGCCAGCTCCCTCA | 10962 | | | |
| | | Db | 181 CCTTAACTATTGGTAGGCAATTAACAAGTCTCCACAGTCACAGCAAGCCAGCTTCTCA | 240 | | | |
| | | OY 10963 | GCCCTGCTTTTATAGAGAGGGGCTCCACGCTCCGATCATGTTTGTGGCCCTCCCTGGA | 11022 | | | |
| | | Db | 241 GCTGTCTTCAATATGAGAGATGTCTCCACCTTCTCATATCTTGTGTGGCTCTC-AGA | 299 | | | |
| | | OY 11023 | CCCGTCCCAAGAGCTCCACATCTTCTCTGATCTGGGGGCCCCAGGCTCGAATGCACTACT | 11082 | | | |
| | | Db | 300 CCTCTCCAAAAAGCTCAATATCTT--CTGTGTGAGTGGGCCCCACAGCTCGATGCACTCA | 357 | | | |
| | | OY 11083 | CCAGATGGGGCCCTCAAAAGACAGTAAAGAGGAGACATACCTTCTCACTCCCTGCTGG | 11142 | | | |
| | | Db | 358 CCAGATGGGGCCCTCAATATGAGTGTGAGAGAGGAGATCTCACTCCCTCACTCAGAGCTGG | 417 | | | |
| | | OY 11143 | CCAGCCCTTCTGTGAGGCCCTGAGTAAACAAGGCTTTC | 11183 | | | |
| | | Db | 418 CCACCTCTTTTGTATGATGACCCAGGAGTACTGTTCTCTGC | 458 | | | |

```

RESULT 11
LOCUS      GGA246112/c      914 bp      DNA      linear      GSS 21-JUN-1999
DEFINITION Gallus gallus anonymous sequence from cosmid mapping to chromosome
6 (Cosmid 56 - Contig 3), genomic survey sequence.
ACCESSION  AJ246112
VERSION     AJ246112.1  GI:5139449
KEYWORDS   GSS: genome survey sequence.
SOURCE     Gallus gallus
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 914)
AUTHORS   Smith,J., Bruley,C.K., Paton,I.R., Law,A.S., Masabanda,J.,
            Maddington,D., Fries,R. and Burt,D.W.
            Integration of the genetic and physical maps of the chicken
            macrochromosomes
TITLE      Unpublished
JOURNAL    2 (bases 1 to 914)
REFERENCE  Law,A.S.
AUTHORS    Direct Submission
TITLE      Submitted (14-JUN-1999) Law A.S., Division of Molecular Biology,
JOURNAL    Roslin Institute, Roslin, Midlothian EH25 9PS, Scotland, UK
FEATURES   location/qualifiers
            source
            1..914
            /organism="Gallus gallus"
            /db_xref="taxon:9031"
            /chromosome="6"
            /map="6q"
BASE COUNT 281 a      179 c      208 g      246 t
ORIGIN
Query Match      1.6%; Score 193.2; DB 17; Length 914;
Best Local Similarity 78.0%; Pred.No.4.8e-23;
Matches 259; Conservative 0; Mismatches 68; Indels 5; Gaps 2;

QY 10533 CTGGGGCTCCAGATTAGTGTAAACCTACTTATACATCATCAATATCATGAAATGGCC 10592
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 643 CTGGAGCTTCATAGAGTGTCAAGTATTCACAACAATCAACAATACAGGAATGGCC 584
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10593 TGGGTGTAAGGAGCCCAAGAGATCATGAAGATCCAAACACCCCGCAGCAGGACCA 10652
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 583 TGGGTGTAAGGAGACCTCAAGATGATGAGCTCCAAACCCCTGCG--CAGCAGGACCA 526
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10653 CCAACCTCCAGATCTGTAGTACAGCAGGACGCCAGGCTCCATCCAACTGGCCATGA 10712
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 525 CCAGCTCCCAT--TTACTAGACCAAGTTGCCAGGCGCTCATCCAACTGGCCTTGA 469
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10713 ACACCTCCAGAGATGAGACATCCAACTCTCTGGGCGCTGTGCGCAGCCTACCA 10772
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 468 ACACCTCCAGAGGAGCGGCGATCCAACTCTCTGGGCGCTGTGCGCAGCCTACCA 409
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10773 CCGTCTGTGTAAGAACTTTCCCTGACATCCATCAATCAAGCTCCCTCGAGGTTAG 10832
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 408 CTCTCTGTGTAAGAACTCTCTCCCTGACATCCAACTCAATCTCCCTCTTCAAGTTAA 349
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10833 ATCCACTCCCTGTGTCTATCAGTCTACT 10864
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 348 AACCGTCCCTGTGTCTGCTTAACACT 317
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
LOCUS      BG624980      437 bp      mRNA      linear      EST 18-APR-2001
DEFINITION pgnlc.pk001.b3 Normalized chicken lymphoid cdna library gallus
gallus cdna clone pgnlc.pk001.b3 5' similar to pif150209|150209
reverse transcriptase VtgIII, pol-I-like - chicken (fragment)
gblAA58720.11 (M28069) unknown protein [Gallus gallus]g, mRNA
sequence.
ACCESSION  BG624980
VERSION     BG624980
FEATURES   sequence.
AUTHORS     Clark,M.S., Edwards,Y.J., McQueen,H.A., Meek,S.E., Smith,S.,
ORIGIN
Query Match      1.6%; Score 191; DB 12; Length 437;
Best Local Similarity 78.2%; Pred.No.1.6e-22;
Matches 241; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 10578 AATCATAGATGGCTGCTGGTTGAAAGGACCCCAAGATCATGAAGATCCAACCCCG 10637
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 307 ACTTATAGATGGCTGTGTTAGAAAGCACCTCAAGATCATGAATCAAAACCCCGN 248
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10638 CCACAGCGGGGCCCAACACCTCCAGATCTGTACTAGACCGGACCCGAGGCTCCAT 10697
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 247 TGTGTCAGGGGCGACCACTTCACATTTATCTAGATCAGCTGCCAGGCTTCAT 188
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10698 CCAACCTGGCGCATGAACACCTCCAGGATGAGATCCACAACTCTCTGGGCGACCTGT 10757
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 187 CCAACCTGGCGCTTGAACAATTGAAGGATGGGCGATCCAAACCTCTCTGGGCGACCTGT 128
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10758 GCGAGCCTCTACACACCTCTCTGTGTAAGAACTTTCCCTGACATCCAACTTAAGCTTC 10817
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 TCCAGACCTCACACACTCTCATAGTAAGTAAGTACTTCCCGCGTATCCAACTTA--TCTTC 69
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10818 CCTCTGTAGTTAGTTCACACCTCCCTGTGCTATCACTGTCTACTGTGTAAGAACT 10877
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 68 CCTCTTCACTTAATAAACCATTTCTCTGTGCTGTGTTATCTGCTTCAAGAGTT 9
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10878 GATTCTCC 10885
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8 GGCTCTCC 1
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
LOCUS      GG0000392/c      608 bp      DNA      linear      GSS 07-SEP-2001
DEFINITION Gallus gallus GSS sequence, clone 901C17bH2, genomic survey
sequence.
ACCESSION  AL605912
VERSION     AL605912.1  GI:15552150
KEYWORDS   GSS: genome survey sequence.
SOURCE     chicken.
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 608)
AUTHORS     Clark,M.S., Edwards,Y.J., McQueen,H.A., Meek,S.E., Smith,S.,

```

```

KEYWORDS   EST.
SOURCE     chicken.
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 437)
AUTHORS   Morgan,R.W. and Burnside,J.
TITLE     Chicken Lymphoid ESTs
JOURNAL   Unpublished (2001)
COMMENT   Contact: Robin W. Morgan
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1341
            Fax: 302 831-2822
            Email: morgan@udel.edu, www.chickest.udel.edu.

FEATURES   location/qualifiers
            source
            1..437
            /organism="Gallus gallus"
            /db_xref="taxon:9031"
            /clone="pgnlc.pk001.b3"
            /clone_lib="Normalized chicken lymphoid cdna library"
            /sex="Male and Female"
            /tissue_type="thymus, bursa, spleen, PBL, bone marrow"
            /lab_host="E.coli EMDH10B"
            /note="Vector: pCMVSPORT 6"
BASE COUNT 110 a      80 c      119 g      107 t      21 others
ORIGIN
Query Match      1.6%; Score 191; DB 12; Length 437;
Best Local Similarity 78.2%; Pred.No.1.6e-22;
Matches 241; Conservative 0; Mismatches 66; Indels 1; Gaps 1;

QY 10578 AATCATAGATGGCTGCTGGTTGAAAGGACCCCAAGATCATGAAGATCCAACCCCG 10637
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 307 ACTTATAGATGGCTGTGTTAGAAAGCACCTCAAGATCATGAATCAAAACCCCGN 248
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10638 CCACAGCGGGGCCCAACACCTCCAGATCTGTACTAGACCGGACCCGAGGCTCCAT 10697
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 247 TGTGTCAGGGGCGACCACTTCACATTTATCTAGATCAGCTGCCAGGCTTCAT 188
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10698 CCAACCTGGCGCATGAACACCTCCAGGATGAGATCCACAACTCTCTGGGCGACCTGT 10757
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 187 CCAACCTGGCGCTTGAACAATTGAAGGATGGGCGATCCAAACCTCTCTGGGCGACCTGT 128
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10758 GCGAGCCTCTACACACCTCTCTGTGTAAGAACTTTCCCTGACATCCAACTTAAGCTTC 10817
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 TCCAGACCTCACACACTCTCATAGTAAGTAAGTACTTCCCGCGTATCCAACTTA--TCTTC 69
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10818 CCTCTGTAGTTAGTTCACACCTCCCTGTGCTATCACTGTCTACTGTGTAAGAACT 10877
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 68 CCTCTTCACTTAATAAACCATTTCTCTGTGCTGTGTTATCTGCTTCAAGAGTT 9
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10878 GATTCTCC 10885
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8 GGCTCTCC 1
            ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
LOCUS      GG0000392/c      608 bp      DNA      linear      GSS 07-SEP-2001
DEFINITION Gallus gallus GSS sequence, clone 901C17bH2, genomic survey
sequence.
ACCESSION  AL605912
VERSION     AL605912.1  GI:15552150
KEYWORDS   GSS: genome survey sequence.
SOURCE     chicken.
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 608)
AUTHORS     Clark,M.S., Edwards,Y.J., McQueen,H.A., Meek,S.E., Smith,S.,

```


| | |
|-----------|---|
| REFERENCE | 1 (bases 1 to 581) |
| AUTHORS | Porter, T.E. and Cogburn, L.A. |
| TITLE | ESTs from Primary Chicken Pituitary/Hypothalamus/Pineal |
| JOURNAL | library USDA/IRAFs Animal Genome Project |
| COMMENT | unpublished (2001) |
| | Contact: Larry A. Cogburn |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 03:56:02 ; Search time 1658 Seconds
(without alignments)
16224.439 Million cell updates/sec

Title: US-09-922-549B-67
Perfect score: 11945
Sequence: 1 tgcgcctctcttgatattc.....tgcgcctctcttgatagg 11945

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Listing first 45 summaries

Database : N.Geneseq_101002:*

- 1: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID82/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID82/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 142 | 1.2 | 22268 | ABQ88219 | Human osteoblast d |
| 2 | 129.4 | 1.1 | 23653 | ABK14793 | Platelet-derived g |
| 3 | 87.2 | 0.7 | 792 | AAO36054 | Human nov gene p |
| 4 | 86 | 0.7 | 385 | ABN79468 | Human ORF4415 cDNA |
| 5 | 85 | 0.7 | 6644 | AAK33181 | Base sequence of t |
| 6 | 85 | 0.7 | 7372 | AAK33182 | Base sequence of t |
| 7 | 85 | 0.7 | 7797 | AAK33180 | Cowpox virus bsr t |
| 8 | 85 | 0.7 | 7996 | AAK33184 | Base sequence of t |
| 9 | 83.4 | 0.7 | 6668 | ABL33697 | Human immune syste |

| | | | | | |
|----|------|-----|-------|----------|--------------------|
| 10 | 79 | 0.7 | 14006 | ABL33958 | Human immune syste |
| 11 | 78.6 | 0.7 | 222 | AAZ98517 | Chicken egg white |
| 12 | 78.2 | 0.7 | 583 | AAO63457 | Chicken (egg white |
| 13 | 77.6 | 0.6 | 97 | AAE67697 | Insulator plasmid |
| 14 | 75.6 | 0.6 | 7442 | AA54668 | Tumour suppressor |
| 15 | 74.6 | 0.6 | 6216 | AAV01425 | Plasmid pRC/ASK8-S |
| 16 | 74.2 | 0.6 | 6024 | ABL33647 | Human immune syste |
| 17 | 74.2 | 0.6 | 9539 | AA545347 | Chemically pretrea |
| 18 | 74.2 | 0.6 | 6292 | ABK28180 | DNA transcription |
| 19 | 73 | 0.6 | 6292 | AA546735 | Tumour suppressor |
| 20 | 72 | 0.6 | 6419 | ABL33267 | Human immune syste |
| 21 | 71.4 | 0.6 | 6898 | ABN80222 | Human chemically m |
| 22 | 71.4 | 0.6 | 7814 | AA546530 | Tumour suppressor |
| 23 | 71 | 0.6 | 6668 | ABL33696 | Human immune syste |
| 24 | 71 | 0.6 | 8298 | ABN80191 | Human immune syste |
| 25 | 70.4 | 0.6 | 7571 | ABL33527 | Human chemically m |
| 26 | 70.4 | 0.6 | 12237 | ABL33458 | Human immune syste |
| 27 | 70 | 0.6 | 15548 | ABL34155 | Human immune syste |
| 28 | 69.8 | 0.6 | 34769 | AA546774 | Human immune syste |
| 29 | 69.6 | 0.6 | 6767 | AA546608 | Tumour suppressor |
| 30 | 69.2 | 0.6 | 16033 | ABL33404 | Human immune syste |
| 31 | 68.4 | 0.6 | 7597 | ABL33013 | Human immune syste |
| 32 | 68 | 0.6 | 5986 | AA545438 | Chemically pretrea |
| 33 | 68 | 0.6 | 6048 | ABO67002 | DNA transcription |
| 34 | 68 | 0.6 | 6048 | ABO67002 | Human angioensis |
| 35 | 68 | 0.6 | 21537 | ABL33999 | Human immune syste |
| 36 | 67.8 | 0.6 | 6030 | AA545456 | Chemically pretrea |
| 37 | 67.8 | 0.6 | 6030 | AA561348 | Human gene regulat |
| 38 | 67.8 | 0.6 | 6030 | ABK28312 | DNA transcription |
| 39 | 67.8 | 0.6 | 6352 | ABL33905 | Human immune syste |
| 40 | 67.8 | 0.6 | 8079 | ABL92313 | Chemically treated |
| 41 | 67.4 | 0.6 | 6145 | ABL32972 | Human immune syste |
| 42 | 67.2 | 0.6 | 887 | AA194064 | Human neuroblastom |
| 43 | 67 | 0.6 | 863 | AA195036 | Human neuroblastom |
| 44 | 66.6 | 0.6 | 866 | AA194008 | Human neuroblastom |
| 45 | 65.8 | 0.6 | 822 | ABQ44852 | Oligonucleotide fo |

ALIGNMENTS

| | |
|------------|---|
| RESULT 1 | |
| ABQ88219/c | |
| ID | ABQ88219 standard; cDNA: 22268 BP. |
| XX | |
| AC | ABQ88219; |
| XX | |
| DT | 18-SEP-2002 (first entry) |
| XX | |
| DE | Human osteoblast differentiation related cDNA seq ID NO 126. |
| XX | |
| KW | Human; osteoblast; stem cell differentiation; bone tissue deposition; |
| KW | osteoporosis; osteopathic; ss. |
| OS | Homo sapiens. |
| XX | |
| PN | W0200250301-A2. |
| PD | 27-JUN-2002. |
| XX | |
| PF | 18-DEC-2001; 2001WO-US48276. |
| XX | |
| PR | 18-DEC-2000; 2000US-255882P. |
| XX | |
| XX | 24-APR-2001; 2001US-285691P. |
| PA | (GENE-) GENE LOGIC INC. |
| PA | (PROC) PROCTER & GAMBLE CO. |
| PI | Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A; |
| PI | Mertz L; |
| XX | |
| DR | WPI; 2002-557663/59. |

PT Use of genes and their expression profiles associated with osteoblast
 PT differentiation for screening modulators bone formation, for diagnosing
 PT or treating e.g. osteoporosis, or as markers for the differentiation
 PT process

PS Claim 1: SEQ ID NO 126; 78pp + Sequence Listing; English.

CC The invention relates to genes and their expression profiles are used
 CC for:

CC (a) screening modulators of precursor stem cell differentiation into
 CC osteoblasts, or bone tissue deposition;

CC (b) diagnosing abnormal deposition of bone tissue, abnormal rate of
 CC osteoblast formation or osteoporosis; or

CC (c) treating or monitoring treatment of the conditions cited in (b), or
 CC monitoring the progression of bone tissue deposition.

CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
 CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy,

CC drug-induced abnormalities in bone formation or bone loss, conditions
 CC that involve altered bone metabolism (e.g. idiopathic juvenile

CC osteoporosis), skeletal disease linked to breast cancer, mastocytosis,
 CC Fanconi syndrome or fibrous dysplasia. The present sequence is that of an

CC osteoblast differentiation associated cDNA marker of the invention.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 22268 BP: 6576 A; 4128 C; 4545 G; 7019 T; 0 other:

Query Match 1.2%; Score 142; DB 24; Length 22268;

Best Local Similarity 68.6%; Pred. No. 1.1e-22;

Matches 232; Conservative 0; Mismatches 90; Indels 16; Gaps 2;

OY 10543 AGATAGGAGTAAACCTTATACATCATCATATCATATGATGGCTGGTTGCA 10602

DB 11362 AGCTACACATTAACTGAGCTATTATTTACTGACATAGAAATGGCTGGTTGAA 11303

OY 10603 GGGACCCCAAGGATGATGAGATGACACACCCCGCACAGGAGGCGCCACCACTTCA 10662

DB 11302 AGGACACACATGATATCTAGTTTCAACCCCGCTGATGATGATGTTGTGACACC- 11244

OY 10663 GATCTGTATAGACAGGAGGAGGCTGATCAACCTGGCCATGACACCTTCAG 10722

DB 11243 -----AGACCAGGCTGCCAGAGCCACATCCAGCTGGCTTGAATGGCTCCAG 11195

OY 10723 GGATGAGATCCCAACATCTCTGCGCAGCTGTGCCAGCCTCACCCTCTCTGT 10782

DB 11194 TGATGAGATCCCAACATCTCTGCGCAGCTGTGCCAGCCTCACCCTCTCTGT 11139

OY 10783 GAAGAATCTTCCCTGACATCCCAATTAAGCTTCCCTGAGGATGATGATCCCTCC 10842

DB 11138 GAAGAATCTTCCCTGACATCCCAATTAAGCTTCCCTGAGGATGATGATCCCTCC 11079

OY 10843 CCTGTGCTATCATGCTGCTACTCTGTAAAGATTGAT 10880

DB 11078 CCTGTGCTATCATGCTGCTACTCTGTAAAGATTGAT 11041

RESULT 2

ABK14793
 ID ABK14793 standard; DNA; 23653 BP.

XX ABK14793;

DT 08-MAY-2002 (first entry)

DE Platelet-derived growth factor expression cassette insert used in pBC734.

XX Platelet-derived growth factor: PDGF, transgenic; milk; gene;

KW wound healing; diabetic foot ulcer; decubitus ulcer; ophthalmic disease;

KW venous stasis ulcer; periodontal regeneration; bone formation;

KW prosthetic vascular graft; pBC734; ds.

XX OS Homo sapiens.

XX WO200198520-A1.

XX 27-DEC-2001.

XX 19-JUN-2001; 2001WO-US41044.

XX 19-JUN-2000; 2000US-212406P.

XX (GENZ) GENZYME TRANSGENICS CORP.

XX Echelard Y, Meade H, Eichner W, Sommermeier K;

XX WPI; 2002-083329/11.

XX Production of platelet derived growth factor (PDGF) comprises

XX expression in the milk of a non-human transgenic animal -

XX Example 1: Fig 1; 59pp; English.

XX The invention relates to the production of platelet-derived growth factor

XX (PDGF) comprising a transgenic mammal whose somatic and germ cells

XX comprise a nucleic acid sequence (I) encoding PDGF, operably linked to a

XX promoter directing expression into mammalian gland epithelial cells, and

XX obtaining the milk from the transgenic mammal where at least 30% of the

XX PDGF in the milk is as a dimer. Also described is a method of producing

XX a transgenic mammal capable of expressing an active PDGF molecule in

XX milk. Pharmaceutical compositions can be obtained from this milk and can

XX be used to stimulate or enhance the wound healing process, in particular

XX diabetic foot ulcers, decubitus ulcers and venous stasis ulcers.

XX Transgenic PDGF (II) can also be used in the treatment of periodontal

XX regeneration, stimulation of bone formation, ophthalmic diseases or

XX healing of prosthetic vascular grafts. (II) can also be used for non-

XX medical applications, e.g., as a supplement for cell culture media or as

XX a component of diagnostic kits. The present sequence represents the

XX platelet-derived growth factor expression cassette insert used in pBC734

XX vector of the invention.

XX Sequence 23653 BP: 6803 A; 5184 C; 5014 G; 6650 T; 2 other:

Query Match 1.1%; Score 129.4; DB 24; Length 23653;

Best Local Similarity 67.7%; Pred. No. 1.2e-19;

Matches 231; Conservative 0; Mismatches 96; Indels 14; Gaps 3;

OY 10565 TATACATATCATAGATCATAGATGAGCTGTTGGAAGGAGCCCAAGATCATGAGA 10624

DB 1470 TTTAGATATGACAGATCATAGATGAGAGCGCTGGTTGCAAGGAGACAGTGTCTATCCAGA 1529

OY 10625 TCCACACCCCCCGGACAGGAGGCGCCACCACTCCAGATGTTGTTACTAGACGAGCGAG 10684

DB 1530 TCCACAC-CCCTGCTATGTGTGAGGGTTCATCAACACGACAGC-----CCAGGCTG 1576

OY 10685 CCCAGGCTCATCCACACTGCGCATGGAACACTCCAGGATGAGATCCACACCTCT 10744

DB 1577 CCCAGGCTCATCCACACTGCGCATGGAACACTCCAGGATGAGATCCACACCTCT 1636

OY 10745 CTGGGAGCTGTGCGACGACCTCCACACCTCTCTGTGAAGAACTTTCCCTGACATCC 10804

DB 1637 TTGGGCAACCTGT-TCAGAGCTGTACACACCTCTGGGGGAAAAAAGCTGCTCTATATCC 1695

OY 10805 AATCAAGCTTCCCTCTCTGAGGTTAGATCCACCTCCCTGTTGCTATCATGCTACT 10864

DB 1696 AACCCAAACCTCCCTGCTGTCTAGTGAATTAAGCCATTCCCTCTGTCTATCAAGGGGAGT 1755

OY 10865 CTGTGAAGAAAGTTGATTCCTCTCTTTTGAAGGTTGCA 10905

DB 1756 TTGCTGTGACATGTTGGTGTGGGTGACACATGTTTGCA 1796

RESULT 3

AAQ36054/C
 ID AAQ36054 standard; cDNA; 792 BP.

XX

AC AA036054;
 XX
 DT 24-MAY-1993 (first entry)
 XX
 DE Chicken nov gene promoter sequence XXXV.
 XX
 XX avian nephroblastoma; avian myeloblastoma virus;
 KW stringent hybridisation; ss.
 XX
 OS Gallus domesticus.
 XX
 PN WO9300430-A.
 XX
 PD 07-JAN-1993.
 XX
 PF 25-JUN-1992; 92MO-FR00589.
 XX
 PR 25-JUN-1991; 91FR-0007807.
 XX
 PA (CNRS) CENT NAT RECH SCI.
 XX
 PI Martinerie C, Perbal B;
 XX
 DR WPI; 1993-036377/04.
 XX
 PT Nucleotide sequences hybridising to regions of chicken nov gene -
 PT useful as probes for detecting complementary sequences to
 PT evaluate development and/or differentiation of tumours
 XX
 PS Disclosure: Page 49-50; 67pp; French.
 XX
 CC The invention includes nucleotide sequences which can hybridise to
 CC all or part of the chicken nov gene under stringent conditions
 CC (i.e. 50% formamide, 5 x SSC). The promoter sequence from the
 CC chicken nov gene, contained in a 1kb SmaI-XhoI fragment and
 CC comprising consensus sequences for different transcription factors
 CC such as Ap2 and Sp1 is specifically disclosed. It also comprises
 CC NF kappa B2 and factor NFkB binding sites.
 XX
 SQ Sequence 792 BP; 158 A; 188 C; 285 G; 161 T; 0 other;
 XX

Query Match 0.7%; Score 87.2; DB 14; Length 792;
 Best Local Similarity 57.8%; Pred. No. 2.4e-10;
 Matches 259; Conservative 0; Mismatches 168; Indels 21; Gaps 5;

QY 10650 CCACCAACCTCCAGATCTGCTACTAGACGACGACGCGGCTCATCCAACTTGCCCA 10709
 DB 440 CAACCCATCTCCTACATTTCTCACTAACAACATCCCTCAGTGCACATCTACACGTTTCC 381
 QY 10710 TGAACACCTCCAGGATGAGATGAGATCCCAACCTCTCTGGGACAGCTGTGCCACACCTCA 10769
 DB 380 TGAACATCTCCAGGATGAGATGATCC--ACCTCGCTGGGACAGATCTGACATCTACTCA 324
 QY 10770 CCACCTCTCTGTG--AAGACCTTTCCCTGACATCCATCTAGCCCTTCCCTCTTGAG 10827
 DB 323 CCATTTTGGGAGAAAAAATTTGTTCTGATATCCAACTGAAATCCCTCCGATTA 264
 QY 10828 GTTAGATCCACCTCCCTTGTGCTATCAGTGTACTCTTGAAGTTGA-----10879
 DB 263 TTTAACACGATTTCTCTCTATCTATCAGTGTACTGAGGAGAGAGGCCGACCTCCACC 204
 QY 10880 --TTTCTCTCTTTTGAAGGTTGCAATGAGCTCTCTGAGCCCTTCTCTCTCTGCG 10937
 DB 203 TCACCAACACCTCTCCAGGAGATTTTAAAGACAATAATGTTTCCCTGAGCCCTCTCC 144
 QY 10938 AGATATGAACAGCCAGCTCCCTCAGCCTGTCTTTATAGAGAGGCTCTCCAGCCCTCTG 10997
 DB 143 GGACTGAACATCCCACTCTCAGCTGCTTCCAAATAAACAATGCTCCACACTCTC 84
 QY 10998 ATATGTTTGTGCTCTCTGACCGCTCCAAAGAGCTCCCAATCTTCTGTTACTGG 11057
 DB 83 ACCACC--TCGTCCCTTCTCTGTGACACACTCCATGGCCCTCGA-----TGTCTTGTAGTGA 30

QY 11058 GGGCCCCAGGCTGAATGCAGTACTCCA 11085
 DB 29 GGGGCCCAAACTGAAACACAGCAGCTCGA 2

RESULT 4
 ABN79468/C
 ID ABN79468 standard; cDNA: 385 BP.
 XX
 AC ABN79468;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE Human ORF4415 cDNA, SEQ ID NO:8829.
 XX
 KW Human; ORF: open reading frame; ORFX: drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cyostatic; nootropic;
 KW neuroprotective; antithrombotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shlunkers RA;
 XX
 DR WPI; 2002-106200/14.
 XX
 PT P-PSDB; ABP35442.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 PS Claim 1; Page 2450; 2508pp; English.

Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,

[illegible]

| | | | |
|------------|--|---|------|
| QY | 346 | TTTTTGAAACGTGACACAGCCCTTTTCTTCAATCCCTTTTGCGTTCTGGCAATGCCCT | 405 |
| Db | 5006 | TT | 4947 |
| QY | 406 | TGCTTTCGATTCGATTAATGAAAGAAACGTGATCGAAGCTTGAGTTTATTTAAGTGG | 465 |
| Db | 4946 | TT | 4887 |
| QY | 466 | GCTTGAAGCTTCGATACGCTGTTGTTACACAGAGATACCTTATTAACTTAAAGCCAGCTTG | 525 |
| Db | 4886 | TT | 4827 |
| QY | 526 | ATGCTTATTTTTTTTCCCTTGAAGTAGTAGGCTTCCTCGGTTTTTTTTTCCCTTGAAGCTG | 585 |
| Db | 4826 | TT | 4767 |
| QY | 586 | GTGAGCGCTTATGATTTTTCATATGAGATTTTTTACCTGATGATCTATGTCATACCAAT | 645 |
| Db | 4766 | TT | 4707 |
| QY | 646 | GCTTGTAAATCTTTTCCCTACTAATACATGTTGATACCTTCGATTTACATGTTGTAATAC | 705 |
| Db | 4706 | TT | 4647 |
| QY | 706 | TTTGCACTCGTGTTTCACGTAATAATATATGATTTTATAGAAATACGTAATCCTGATT | 765 |
| Db | 4646 | TT | 4587 |
| QY | 766 | TCCCTTTTATTCATCTATATGCTCTGTGTTACAGTCAACAGACTCCACTCTATT | 825 |
| Db | 4586 | TT | 4527 |
| QY | 826 | TATTTATAGAAATTTTATATGACAGCTCTGCTGGTTCTTGCTGTAGAGATACAGCCT | 885 |
| Db | 4526 | TT | 4467 |
| QY | 886 | AAATTTCCTAGAGGAGATGCTCAGTAAAGCGCGGT | 918 |
| Db | 4466 | GCAACTGCAAGAGGGTTATTGATACACAGGGT | 4434 |
| RESULT 7 | | | |
| AAK33180/c | | | |
| ID | AAK33180 standard; DNA; 7797 BP. | | |
| XX | AAK33180; | | |
| XX | 25-JUN-1999 (first entry) | | |
| XX | Cowpox virus bsr full length gene sequence. | | |
| XX | Cowpox virus. | | |
| XX | WO913073-A2. | | |
| XX | 18-MAR-1999. | | |
| XX | 07-SEP-1998; 98WO-JP04010. | | |
| XX | 08-SEP-1997; 97JP-0259235. | | |
| XX | (RPRG-) RPR GENCELL ASIA PACIFIC INC. | | |
| XX | Hamada H; | | |
| XX | WPI; 1999-243728/20. | | |
| XX | New apoptosis-resistant virus-sensitive cell | | |

| | | | |
|----------|----------|--|------|
| OY | 286 | AAATTTCTACTGACGTGATGTTTGGACAGGTACAGAAACATTTCTCAAAAGAAAGAAC | 345 |
| | | | |
| Db | 5690 | TT | 5631 |
| OY | 346 | TTTTGGAAACGTACACACCCTTTTCCTTCATTCCTCTTTGCTTCTGCGCAATGCTT | 405 |
| | | | |
| Db | 5630 | TT | 5571 |
| OY | 406 | TGGTTCGATGTCATTAATGAGAAACGTGATCGCAGACTGAGTTTATATATATGCTG | 465 |
| | | | |
| Db | 5570 | TT | 5511 |
| OY | 466 | GCTTGAAAGCTGAGATGAGCTGTGTACACAGATACCTATTAACTTATAGGCGACCTG | 525 |
| | | | |
| Db | 5510 | TT | 5451 |
| OY | 526 | ATGCTTATTTTTTCCCTTGAAGTACGAGCTCTCGGTTTTTTCCTTGAACCTG | 585 |
| | | | |
| Db | 5450 | TTATTT | 5391 |
| OY | 586 | GTGAGCCTTACATTTTTCTATGAGGATTTTTTACCTGATGATCTAGTTGATCCCAAT | 645 |
| | | | |
| Db | 5390 | TT | 5331 |
| OY | 646 | GCTGTGAAAGTTTTCCTAGTACATGTTGATTAACCTCGGATTTACATGTTGTATAC | 705 |
| | | | |
| Db | 5330 | TT | 5271 |
| OY | 706 | TTGCATCTGTGCTCTAGAGAAATATATGACATTTATAGAAATACGTAATTCCTGAT | 765 |
| | | | |
| Db | 5270 | TT | 5211 |
| OY | 766 | TCTTTTTTTTTTATCTCTATGCTCTGTGTACAGGTCAAGACTTCACCTCATTTT | 825 |
| | | | |
| Db | 5210 | TT | 5151 |
| OY | 826 | TATTTATAGAAATTTATATGACAGCTGTGCTTGCTTCTGTGTGATAGATACAGCCT | 885 |
| | | | |
| Db | 5150 | TT | 5091 |
| OY | 886 | AAATTTCTAGAGCATGCTCAGTAAGCGGGT | 918 |
| | | | |
| Db | 5090 | GCAACTGCAGAGGGTTTATTGATACACGGT | 5058 |
| RESULT 9 | | | |
| ID | ABL33697 | standard; DNA: 6668 BP. | |
| XX | XX | ABL33697: | |
| AC | XX | ABL33697: | |
| DT | XX | 26-MAR-2002 (first entry) | |
| XX | XX | | |
| DE | XX | Human immune system associated gene seq ID NO: 1670. | |
| XX | KW | Human: immune system disease; cytosine methylation; antiasthmatic; | |
| XX | KW | antileptoterosclerotic; anti-leukemia; cytosolic; nocturnal; | |
| XX | KW | neuroprotective; anti-HIV; anticonvulsant; ophthalmological; | |
| XX | KW | antiinflammatory; antiarthritic; antidiabetic; antipsoriatic; | |
| XX | KW | antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia; | |
| KW | KW | acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy; | |
| KW | KW | neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; | |
| XX | XX | gene; ds. | |
| OS | XX | Homo sapiens. | |
| XX | PN | WO200200928-A2. | |
| XX | XX | | |
| PD | XX | 03-JAN-2002. | |
| XX | XX | | |
| PF | XX | 02-JUL-2001; 2001WO-EP07537. | |

| DB | 3235 | 3236 | 3237 | 3238 | 3239 | 3240 | 3241 | 3242 | 3243 | 3244 | 3245 | 3246 | 3247 | 3248 | 3249 | 3250 | 3251 | 3252 | 3253 | 3254 | 3255 | 3256 | 3257 | 3258 | 3259 | 3260 | 3261 | 3262 | 3263 | 3264 | 3265 | 3266 | 3267 | 3268 | 3269 | 3270 | 3271 | 3272 | 3273 | 3274 | 3275 | 3276 | 3277 | 3278 | 3279 | 3280 | 3281 | 3282 | 3283 | 3284 | 3285 | 3286 | 3287 | 3288 | 3289 | 3290 | 3291 | 3292 | 3293 | 3294 | 3295 | 3296 | 3297 | 3298 | 3299 | 3300 | 3301 | 3302 | 3303 | 3304 | 3305 | 3306 | 3307 | 3308 | 3309 | 3310 | 3311 | 3312 | 3313 | 3314 | 3315 | 3316 | 3317 | 3318 | 3319 | 3320 | 3321 | 3322 | 3323 | 3324 | 3325 | 3326 | 3327 | 3328 | 3329 | 3330 | 3331 | 3332 | 3333 | 3334 | 3335 | 3336 | 3337 | 3338 | 3339 | 3340 | 3341 | 3342 | 3343 | 3344 | 3345 | 3346 | 3347 | 3348 | 3349 | 3350 | 3351 | 3352 | 3353 | 3354 | 3355 | 3356 | 3357 | 3358 | 3359 | 3360 | 3361 | 3362 | 3363 | 3364 | 3365 | 3366 | 3367 | 3368 | 3369 | 3370 | 3371 | 3372 | 3373 | 3374 | 3375 | 3376 | 3377 | 3378 | 3379 | 3380 | 3381 | 3382 | 3383 | 3384 | 3385 | 3386 | 3387 | 3388 | 3389 | 3390 | 3391 | 3392 | 3393 | 3394 | 3395 | 3396 | 3397 | 3398 | 3399 | 3400 | 3401 | 3402 | 3403 | 3404 | 3405 | 3406 | 3407 | 3408 | 3409 | 3410 | 3411 | 3412 | 3413 | 3414 | 3415 | 3416 | 3417 | 3418 | 3419 | 3420 | 3421 | 3422 | 3423 | 3424 | 3425 | 3426 | 3427 | 3428 | 3429 | 3430 | 3431 | 3432 | 3433 | 3434 | 3435 | 3436 | 3437 | 3438 | 3439 | 3440 | 3441 | 3442 | 3443 | 3444 | 3445 | 3446 | 3447 | 3448 | 3449 | 3450 | 3451 | 3452 | 3453 | 3454 | 3455 | 3456 | 3457 | 3458 | 3459 | 3460 | 3461 | 3462 | 3463 | 3464 | 3465 | 3466 | 3467 | 3468 | 3469 | 3470 | 3471 | 3472 | 3473 | 3474 | 3475 | 3476 | 3477 | 3478 | 3479 | 3480 | 3481 | 3482 | 3483 | 3484 | 3485 | 3486 | 3487 | 3488 | 3489 | 3490 | 3491 | 3492 | 3493 | 3494 | 3495 | 3496 | 3497 | 3498 | 3499 | 3500 | 3501 | 3502 | 3503 | 3504 | 3505 | 3506 | 3507 | 3508 | 3509 | 3510 | 3511 | 3512 | 3513 | 3514 | 3515 | 3516 | 3517 | 3518 | 3519 | 3520 | 3521 | 3522 | 3523 | 3524 | 3525 | 3526 | 3527 | 3528 | 3529 | 3530 | 3531 | 3532 | 3533 | 3534 | 3535 | 3536 | 3537 | 3538 | 3539 | 3540 | 3541 | 3542 | 3543 | 3544 | 3545 | 3546 | 3547 | 3548 | 3549 | 3550 | 3551 | 3552 | 3553 | 3554 | 3555 | 3556 | 3557 | 3558 | 3559 | 3560 | 3561 | 3562 | 3563 | 3564 | 3565 | 3566 | 3567 | 3568 | 3569 | 3570 | 3571 | 3572 | 3573 | 3574 | 3575 | 3576 | 3577 | 3578 | 3579 | 3580 | 3581 | 3582 | 3583 | 3584 | 3585 | 3586 | 3587 | 3588 | 3589 | 3590 | 3591 | 3592 | 3593 | 3594 | 3595 | 3596 | 3597 | 3598 | 3599 | 3600 | 3601 | 3602 | 3603 | 3604 | 3605 | 3606 | 3607 | 3608 | 3609 | 3610 | 3611 | 3612 | 3613 | 3614 | 3615 | 3616 | 3617 | 3618 | 3619 | 3620 | 3621 | 3622 | 3623 | 3624 | 3625 | 3626 | 3627 | 3628 | 3629 | 3630 | 3631 | 3632 | 3633 | 3634 | 3635 | 3636 | 3637 | 3638 | 3639 | 3640 | 3641 | 3642 | 3643 | 3644 | 3645 | 3646 | 3647 | 3648 | 3649 | 3650 | 3651 | 3652 | 3653 | 3654 | 3655 | 3656 | 3657 | 3658 | 3659 | 3660 | 3661 | 3662 | 3663 | 3664 | 3665 | 3666 | 3667 | 3668 | 3669 | 3670 | 3671 | 3672 | 3673 | 3674 | 3675 | 3676 | 3677 | 3678 | 3679 | 3680 | 3681 | 3682 | 3683 | 3684 | 3685 | 3686 | 3687 |
|----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|

CC preservation of DNA on the support and allows simple distribution of the
CC supported DNA through the mail. The sheets can be stored at normal
CC temperatures, eliminating the need for low temperature storage of frozen
CC aqueous solutions or glycerol-containing stock solutions. The DNA
CC solution may be applied using a standard ink-jet printer, simply by
CC replacing the ink cartridge with one containing the DNA solution. Very
CC many DNAs, e.g. 1000, can be attached to the same support so that a
CC booklet of 100 sheets will contain the complete set of murine cDNAs.
CC The present sequence represents a chicken egg white lysozyme cDNA
CC fragment. This is used in demonstrating a typical support sheet
CC carrying lysozyme cDNA as a series of dots, which may be cut out
CC individually. The sheet also contains sequence and structural information
CC for lysozyme.

SQ Sequence 222 BP; 44 A; 50 C; 65 G; 63 T; 0 other;

Query Match 0.7%; Score 78.6; DB 21; Length 222;
Best Local Similarity 95.3%; Pred. No. 1.4e-08;
Matches 81; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 11861 GCAGTCCCGCTGTGTACGACACTGGCAACATGAGGCTTTGCTAATCTGTGCTTTG 11920
DB 1 GCAGTCCCGCTGTGTACGACACTGGCAACATGAGGCTTTGCTAATCTGTGCTTTG 60

OY 11921 CTTCTGCCCTGTGCTGCTTGAAGG 11945
DB 61 CTTCTGCCCTGTGCTGCTGCTGCGG 85

RESULT 12
AA063457 standard; DNA; 583 BP.

XX AA063457;

DT 24-SEP-1994 (first entry)

DE Chicken (egg white) lysozyme gene.

XX Lytic peptide; resistance: fire blight; pomaceous fruit; scion;
KW rootstock cultivar; transformation; transgenic variants; pears;
KM apples; quinces; ss.

OS Chicken.

XX Key Location/Qualifiers

FT CDS 30..470

FT CDS /*tag= a

XX WO9407356-A.

PD 14-APR-1994.

PF 29-SEP-1993; 93WO-US09299.

PR 30-SEP-1992; 92US-0954347.

PR 18-MAR-1993; 93US-0033772.

XX (CORR) CORNELL RES FOUND INC.

PI Aldwinckle HS, Norelli JL;

XX WPI; 1994-135092/16.

DR P-PSDB; AAR53348.

XX Prod'n of transgenic pomaceous fruit with fire blight resistance -

CC by transforming fruit scion or rootstock cultivars with a gene
PS encoding a lytic protein, e.g. lysozyme
XX Disclosure: Page 38C; 61pp; English.
XX Resistance to fire blight conferred on pomaceous fruit scion or
CC rootstock cultivars can be induced by transforming the plant with a

CC gene encoding a lytic protein, e.g. lysozyme, cecropin, attacin or
CC homologues. The transgenic variants show resistance to fire blight
CC but retain their own characteristics. The method can be used to
CC protect species of the Rosaceae family, e.g. pears, apples and
CC quinces.
CC See also AA063455-69.

SQ Sequence 583 BP; 133 A; 168 C; 172 G; 110 T; 0 other;

Query Match 0.7%; Score 78.2; DB 15; Length 583;
Best Local Similarity 96.4%; Pred. No. 2.9e-08;
Matches 80; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 11863 AGTCCCGCTGTGTACGACACTGGCAACATGAGGCTTTGCTAATCTGTGCTTTGCT 11922
DB 1 AGTCCCGCTGTGTACGACACTGGCAACATGAGGCTTTGCTAATCTGTGCTTTGCT 60

OY 11923 TCCTGCCCTGTGCTGCTTGAAGG 11945
DB 61 TCCTGCCCTGTGCTGCTGCTGCGG 83

RESULT 13

AAF67697 standard; DNA; 97 BP.

XX AAF67697;

DT 12-APR-2001 (first entry)

DE Insulator plasmid enhancer blocking sequence lys SEQ ID NO: 55.

XX Chicken; human; insulator; enhancer; DNA binding protein;
KW gene expression; gene therapy; insulin-like growth factor-2; Igf2;
KM knockout mouse; ds.

OS Unidentified.

XX WO200102553-A2.

PD 11-JAN-2001.

PF 19-APR-2000; 2000WO-US10509.

PR 30-JUN-1999; 99US-0141728.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Bell AC, West AG, Felsenfeld G;

PI WPI; 2001-091803/10.

DR Isolated DNA molecule useful for the regulation of gene expression and

XX function in mammals and plants -

PS Example 1; Page 65; 96pp; English.

CC The present invention provides the sequence of a enhancer-blocking
CC insulator from the chicken. Also provided are insulators from the murine,
CC rat and human insulin-like growth factor-2 (Igf2) genes. The insulators
CC can be used to modulate gene expression, for example in gene therapy and
CC in knockout mouse production.

SQ Sequence 97 BP; 22 A; 20 C; 23 G; 32 T; 0 other;

Query Match 0.6%; Score 77.6; DB 22; Length 97;
Best Local Similarity 95.2%; Pred. No. 1.5e-08;
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 9435 GAGGTTTACGCTGATTTGACATGAGAAATGAGACCTCTACTGATAGCTATGCTAT 9494
DB 5 GAGGCTTACGCTGATTTGACATGAGAAATGAGACCTCTACTGATAGCTATGCTAT 64

QY 9495 TACGTCCTTTTGGCTACTACT 9518
||| ||||||||||||||||
Db 65 TACATGCTTTTGCTTAAGTACT 88

RESULT 14
ID AAS46686
AAS46686 standard; DNA; 7442 BP.
XX AC AAS46686;
XX DT 18-DEC-2001 (first entry)
XX DE Tumour suppressor gene derived chemically modified sequence #409.
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
OS Homo sapiens.
XX WO200168912-A2.
PN 20-SEP-2001.
XX 15-MAR-2001; 2001WO-BP02955.
PF 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1033529.
PR 01-SEP-2000; 2000DE-1043826.
PA (EPIC-) EPIGENOMICS AG.
PI Olek A, Piepenbrock C, Berlin K;
DR WPL: 2001-602752/68.

PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX
XX
XX Claim 1; SEQ ID No 409; 27pp; English.
PS
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 7442 BP: 1655 A; 214 C; 1934 G; 3636 T; 3 other;

Query Match 0.6%; Score 75.6; DB 22; Length 7442;

```

XX 13-MAR-1996: 96US-0013321.
XX
XX (SALA-) SOUTH ALABAMA MEDICAL SCI FOUND.
XX
XX PI Kayes SG;
XX
XX WPI: 1997-470644/43.
XX
XX PT Vaccine for schistosomiasis - comprising a non-integrating DNA
XX PT sequence encoding Schistosoma mansoni glutathione S-transferase
XX
XX PS Claim 4: Fig 7a; 57pp; English.
XX
XX CC A vaccine has been developed comprising a non-infectious,
XX CC non-integrating DNA sequence encoding Schistosoma mansoni
XX CC glutathione S-transferase (GST). The DNA comprises a plasmid
XX CC encoding a protein, polypeptide or peptide which is operably
XX CC linked to a promoter. The present sequence represents a
XX CC specifically claimed plasmid, pRC/ASK8-SmGST3. The vaccine
XX CC can be used for preventing or ameliorating the effects of
XX CC schistosomiasis infection. They elicit immunological responses
XX CC necessary for inducing resistance to infection and/or ameliorating
XX CC the disease in such as way as to alter subsequent transmission of
XX CC the disease.
XX
XX SO Sequence 6216 BP; 1428 A; 1624 C; 1627 G; 1537 T; 0 other;

Query Match 0.68; Score 74.6; DB 18; Length 6216;
Best Local Similarity 61.28; Pred. No. 7.6e-07;
Matches 139; Conservative 0; Mismatches 84; Indels 4; Gaps 1;

QY 10573 CATAGATCATAGATGGCTGGGTGGAGGAGGCCCAAGATCATGAGATCCACAC 10632
DB 432 CCTAGTATCATAGATGACAGGTGGAATACTTAAGATCATGAGTCCACCAT 373
QY 10633 CCCCAGCAGCAGGAGCCACCACTCCAGATCTGTAGACGACGAGCCAGGAG 10692
DB 372 GACCTAACCATACCTACTACTAC---AACCTCCGTTAATCATGTCCCGAGCAC 317
QY 10693 TCCATCCAACTGAGCATGACCTCCAGGATGAGACATCCAACTCTCTGGGAG 10752
DB 316 CACATCCAAACGGTTTAAAGCATCATCAGGAGTGACCAACACCTCCCTGGGGAG 257
QY 10753 CCTGTGCCAGCACCCTCACCACCTCTCTGTGAAGAACTTTCCCTGA 10799
DB 256 CCTATTCCAGTGTATTAACACCTTTATGTAAAGATTTCCTGA 210

```

Search completed: March 27, 2003, 04:42:19
 Job time : 2132 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 03:59:18 ; Search time 265 Seconds
(without alignments)
13823.600 Million cell updates/sec

Title: US-09-922-549b-67

Perfect score: 11945
Sequence: 1 tgcgcctcttgatattc.....tgcgcctgctgcttagg 11945

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------------------------|
| 1 | 80.2 | 0.7 | 222 | 4 | US-09-388-917-1 Sequence 1, Appl |
| 2 | 78.2 | 0.7 | 586 | 1 | US-08-385-590A-3 Sequence 3, Appl |
| 3 | 78.2 | 0.7 | 586 | 3 | US-09-021-520-3 Sequence 3, Appl |
| 4 | 62.6 | 0.5 | 7218 | 1 | US-08-232-463-14 Sequence 14, Appl |
| 5 | 56.4 | 0.5 | 1987 | 1 | US-07-999-280A-23 Sequence 23, Appl |
| 6 | 56.4 | 0.5 | 1987 | 1 | US-08-426-036-23 Sequence 23, Appl |
| 7 | 56.4 | 0.5 | 1987 | 1 | US-08-426-279-23 Sequence 23, Appl |
| 8 | 56.4 | 0.5 | 1987 | 1 | US-08-401-013-23 Sequence 23, Appl |
| 9 | 56.4 | 0.5 | 1987 | 3 | US-08-426-570-23 Sequence 23, Appl |
| 10 | 56.4 | 0.5 | 1987 | 3 | US-08-426-876-23 Sequence 23, Appl |
| 11 | 56.4 | 0.5 | 1987 | 3 | US-08-426-243-23 Sequence 23, Appl |
| 12 | 56.4 | 0.5 | 1987 | 4 | US-08-401-632-23 Sequence 23, Appl |
| 13 | 54.6 | 0.5 | 19124 | 2 | US-08-487-826B-13 Sequence 13, Appl |
| 14 | 53.4 | 0.4 | 974 | 3 | US-08-675-885-3 Sequence 3, Appl |
| 15 | 49.2 | 0.4 | 2246 | 4 | US-09-363-708-3 Patent No. 5168051 |
| 16 | 49.2 | 0.4 | 2671 | 6 | US-08-232-463-14 Patent No. 5168051 |
| 17 | 48.8 | 0.4 | 7218 | 4 | US-09-071-224-3 Sequence 3, Appl |
| 18 | 48.6 | 0.4 | 1813 | 4 | US-08-998-416-595 Sequence 55, Appl |
| 19 | 48.2 | 0.4 | 658 | 4 | US-09-465-558-53 Sequence 53, Appl |
| 20 | 48.2 | 0.4 | 1206 | 4 | US-08-897-340-4 Sequence 4, Appl |
| 21 | 48 | 0.4 | 1700 | 2 | US-09-252-329-4 Sequence 6, Appl |
| 22 | 48 | 0.4 | 1700 | 3 | US-08-628-417-6 Sequence 6, Appl |
| 23 | 47.8 | 0.4 | 240 | 1 | US-08-628-417-5 Sequence 5, Appl |
| 24 | 47.4 | 0.4 | 140 | 1 | US-08-462-467B-13 Sequence 13, Appl |
| 25 | 47.4 | 0.4 | 2252 | 4 | US-08-334-179A-13 Sequence 13, Appl |
| 26 | 47.4 | 0.4 | 2252 | 4 | US-09-336-536-66 Sequence 66, Appl |
| 27 | 47.2 | 0.4 | 1937 | 4 | US-09-336-536-66 Sequence 66, Appl |

| | | | | | |
|------|------|-----|------|---|-------------------------------------|
| C 28 | 47 | 0.4 | 2323 | 4 | US-09-149-476-24 Sequence 24, Appl |
| C 29 | 46.4 | 0.4 | 144 | 1 | US-08-702-344-26 Sequence 26, Appl |
| C 30 | 46.4 | 0.4 | 1098 | 3 | US-09-248-335-35 Sequence 35, Appl |
| C 31 | 46.2 | 0.4 | 1023 | 1 | US-08-252-966B-16 Sequence 16, Appl |
| C 32 | 46.2 | 0.4 | 2447 | 2 | US-09-014-969-14 Sequence 14, Appl |
| C 33 | 46 | 0.4 | 578 | 4 | US-09-602-877A-95 Sequence 27, Appl |
| C 34 | 46 | 0.4 | 1447 | 4 | US-09-443-041A-27 Sequence 4, Appl |
| C 35 | 46 | 0.4 | 6755 | 3 | US-08-931-999-4 Sequence 17, Appl |
| C 36 | 45.8 | 0.4 | 1307 | 2 | US-08-960-022-17 Sequence 9, Appl |
| C 37 | 45.6 | 0.4 | 1172 | 1 | US-07-945-288-9 Sequence 9, Appl |
| C 38 | 45.6 | 0.4 | 1172 | 1 | US-08-462-831-9 Sequence 9, Appl |
| C 39 | 45.6 | 0.4 | 1172 | 1 | US-08-461-809-9 Sequence 9, Appl |
| C 40 | 45.6 | 0.4 | 1172 | 1 | US-08-461-441-9 Sequence 9, Appl |
| C 41 | 45.6 | 0.4 | 1172 | 5 | PCR-US93-08518-9 Sequence 9, Appl |
| C 42 | 45.4 | 0.4 | 6004 | 4 | US-08-961-527-27 Sequence 27, Appl |
| C 43 | 45.2 | 0.4 | 1445 | 4 | US-09-814-951A-1 Sequence 1, Appl |
| C 44 | 45 | 0.4 | 2852 | 3 | US-09-027-137-2 Sequence 2, Appl |
| C 45 | 45 | 0.4 | 2852 | 4 | US-09-344-441-2 Sequence 2, Appl |

ALIGNMENTS

RESULT 1
US-09-388-917-1
Sequence 1, Application US/09388917
Patent No. 6258542
GENERAL INFORMATION:
APPLICANT: Hayashizaki, Yoshihide
TITLE OF INVENTION: Method for Supporting DNA-Fixation and DNA-Fixed
FILE REFERENCE: 1794-0120P
CURRENT APPLICATION NUMBER: US/09/388, 917
CURRENT FILING DATE: 1999-09-02
EARLIER APPLICATION NUMBER: 10-250619
EARLIER FILING DATE: 1998-09-04
EARLIER APPLICATION NUMBER: 11-240910
EARLIER FILING DATE: 1999-08-27
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 222
TYPE: DNA
ORGANISM: Gallus gallus
US-09-388-917-1

Query Match 0.7%; Score 80.2; DB 4; Length 222;
Best Local Similarity 96.5%; Pred. No. 1.3e-11;
Matches 82; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11861 GCAGCCCGCGTGTGTGACACACATGAGCTTGTGAATCTGTGCTTTG 11920
|||||
DB 1 GCAGTCCCGCTGTGTGTGACACACTGCGACATGAGCTTGTGAATCTGTGCTTTG 60
|||||

QY 11921 CTTCTGCCCTGGCTGCTAGG 11945
|||||

DB 61 CTTCTGCCCTGGCTGCTAGG 85
|||||

RESULT 2
US-08-385-590A-3
Sequence 3, Application US/08385590A
Patent No. 5824861
GENERAL INFORMATION:
APPLICANT: Alwinckle, Herbert S.
APPLICANT: No. 5824861e11, John I.
TITLE OF INVENTION: TRANSGENIC POMACEOUS FRUIT WITH FIRE
TITLE OF INVENTION: BLIGHT RESISTANCE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square
CITY: Rochester

STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,590A
FILING DATE: 08-FEB-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,347
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/143
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 586 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-385-590A-3

Query Match 0.7%; Score 78.2; DB 1; Length 586;
Best Local Similarity 96.4%; Pred. No. 8,4e-11;
Matches 80; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 11863 AGTCCCGCTGTGTACGACACTGGCAACATGAGCTTTGCTAACTTGTCCTTGGCT 11922
Db 1 AGTCCCGCTGTGTACGACACTGGCAACATGAGCTTTGCTAACTTGTCCTTGGCT 60
Qy 11923 TCCTGCCCTGCTGCTTACGGG 11945
Db 61 TCCTGCCCTGCTGCTTACGGG 83

RESULT 3
US-09-021-520-3
Sequence 3, Application US/09021520
Patent No. 6100453
GENERAL INFORMATION:
APPLICANT: Aldwinckle, Herbert S.
APPLICANT: No. 6100453e11, John L.
TITLE OF INVENTION: TRANSGENIC POMACEOUS FRUIT WITH FIRE
BLIGHT RESISTANCE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,520
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/385,590
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/143
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 586 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-09-021-520-3
Query Match 0.7%; Score 78.2; DB 3; Length 586;
Best Local Similarity 96.4%; Pred. No. 8,4e-11;
Matches 80; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 11863 AGTCCCGCTGTGTACGACACTGGCAACATGAGCTTTGCTAACTTGTCCTTGGCT 11922
Db 1 AGTCCCGCTGTGTACGACACTGGCAACATGAGCTTTGCTAACTTGTCCTTGGCT 60
Qy 11923 TCCTGCCCTGCTGCTTACGGG 11945
Db 61 TCCTGCCCTGCTGCTTACGGG 83

RESULT 4
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 70..1726
NAME/KEY: mat_peptide
LOCATION: 166..1726
US-08-426-036-23

Query Match
Best Local Similarity 59.3%; Pred. No. 0.00013;
Matches 96; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

OY 11037 TCCACATCTTTCCTGCTACTGGGGCCCCAGGCTGATGACGTAAGTGGGGCTC 11096
DB 1772 TCAATGTCCCTCTGAAATGTGGCGCCAGCCCTGGACACAGTACTCCAGATGTGTCTG 1831
OY 11097 AAAGAGCAGAGTAAGAGGAGCAATCACCTTCTCAACCCTGCTGGCCAGCCCTTCTG 11156
DB 1832 ACCAGCTCAGAGTACAGTGGGACGCTGTCTCTTCTGATCTGACAGTACTCTTCTACTC 1891
OY 11157 ATGAGCCCTGGATCAACTGCTTCTGAGCTGCAACTCT 11198
DB 1892 GTGCAGATTAGATCACATTAAGTTTAAACAGCTGCATCATAT 1933

RESULT 7
US-08-426-279-23
Sequence 23, Application US/08426279
Patent No. 5672343
GENERAL INFORMATION:
APPLICANT: LADNER, MARTHA B.
APPLICANT: NOBLE, JANELLE A.
APPLICANT: MARTIN, GEORGE A.
APPLICANT: KAWASAKI, ERNEST S.
APPLICANT: COVNE, MAZIE YEE
APPLICANT: HALENBECK, ROBERT F.
APPLICANT: KOTHS, KIRSTON E.
TITLE OF INVENTION: NEW FORMS OF COLONY STIMULATING FACTOR-1
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: Intellectual Property - R440, P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,279
FILING DATE: 21-APR-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mcgarigle Jr., Phillip L.
REGISTRATION NUMBER: 31,395
REFERENCE/DOCKET NUMBER: 0681.013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2718
TELEFAX: (510) 655-3542
TELEX: n/a
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1987 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 70..1726
NAME/KEY: mat_peptide
LOCATION: 166..1726
US-08-426-279-23

Query Match
Best Local Similarity 59.3%; Pred. No. 0.00013;
Matches 96; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

OY 11037 TCCACATCTTTCCTGCTACTGGGGCCCCAGGCTGATGACGTAAGTGGGGCTC 11096
DB 1772 TCAATGTCCCTCTGAAATGTGGCGCCAGCCCTGGACACAGTACTCCAGATGTGTCTG 1831
OY 11097 AAAGAGCAGAGTAAGAGGAGCAATCACCTTCTCAACCCTGCTGGCCAGCCCTTCTG 11156
DB 1832 ACCAGCTCAGAGTACAGTGGGACGCTGTCTCTTCTGATCTGACAGTACTCTTCTACTC 1891
OY 11157 ATGAGCCCTGGATCAACTGCTTCTGAGCTGCAACTCT 11198
DB 1892 GTGCAGATTAGATCACATTAAGTTTAAACAGCTGCATCATAT 1933

RESULT 8
US-08-401-013-23
Sequence 23, Application US/08401013
Patent No. 5681719
GENERAL INFORMATION:
APPLICANT: LADNER, MARTHA B.
APPLICANT: NOBLE, JANELLE A.
APPLICANT: MARTIN, GEORGE A.
APPLICANT: KAWASAKI, ERNEST S.
APPLICANT: COVNE, MAZIE YEE
APPLICANT: HALENBECK, ROBERT F.
APPLICANT: KOTHS, KIRSTON E.
TITLE OF INVENTION: NEW FORMS OF COLONY STIMULATING FACTOR-1
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: Intellectual Property - R440, P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,013
FILING DATE: 08-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/999,280
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mcgarigle Jr., Phillip L.
REGISTRATION NUMBER: 31,395
REFERENCE/DOCKET NUMBER: 0681.007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2718
TELEFAX: (510) 655-3542
TELEX: n/a
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1987 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 70..1726
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 166..1726
US-08-401-013-23

Query Match
Best Local Similarity 59.3%; Score 56.4; DB 1; Length 1987;
Pred. No. 0.00013;
Matches 96; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 11037 TCACATCTTCTGCTGACTGGGGGCCCCAGGCTGTAATGCAGTACTCCAGATGGGGCTTC 11096

DB 1772 TCATATGCTCCCTGGAATAATGTGGCGCCAGCCCTGGACACAGTACTCCAGATGTTCTCTG 1831

QY 11097 AAAAGAGCAGATAAGAGGAGCAATCACCCTTCCTCAGCCCTGGCCAGCCCTCTCTG 11156

DB 1832 ACCAGCTCAGAGTACAGTGGGAGCGTTGCTTCTTGATCTGGACAGTACTCTTCTACTC 1891

QY 11157 ATGAGCCCTGGATACACTGGCTTCTGAGCTGCACTTCT 11198

DB 1892 GTGCAGATTAGATCATTAGTTTAAACAGCTGCATCATAT 1933

RESULT 9

US-08-426-570-23

Sequence 23, Application US/08426570

Patent No. 6103224

GENERAL INFORMATION:

APPLICANT: LADNER, MARTHA B.

APPLICANT: NOBLE, JANELLE A.

APPLICANT: MARTIN, GEORGE A.

APPLICANT: KAWASAKI, ERNEST S.

APPLICANT: COYNE, MAZIE YEE

APPLICANT: HALENECK, ROBERT F.

APPLICANT: KOTHS, KIRSTON E.

TITLE OF INVENTION: NEW FORMS OF COLONY STIMULATING FACTOR-1

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHIRON CORPORATION

STREET: Intellectual Property - R440, P.O. Box 8097

CITY: Emeryville

STATE: California

COUNTRY: U.S.A.

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/426,570

FILING DATE: 21-APR-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: McGarrigle Jr., Philip L.

REGISTRATION NUMBER: 31,395

REFERENCE/DOCKET NUMBER: 0681,012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 601-2718

TELEFAX: (510) 655-3542

TELEX: n/a

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1987 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 70..1726
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 166..1726
US-08-426-570-23

Query Match
Best Local Similarity 59.3%; Score 56.4; DB 3; Length 1987;
Pred. No. 0.00013;
Matches 96; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 11037 TCACATCTTCTGCTGACTGGGGGCCCCAGGCTGTAATGCAGTACTCCAGATGGGGCTTC 11096

DB 1772 TCATATGCTCCCTGGAATAATGTGGCGCCAGCCCTGGACACAGTACTCCAGATGTTCTCTG 1831

QY 11097 AAAAGAGCAGATAAGAGGAGCAATCACCCTTCCTCAGCCCTGGCCAGCCCTCTCTG 11156

DB 1832 ACCAGCTCAGAGTACAGTGGGAGCGTTGCTTCTTGATCTGGACAGTACTCTTCTACTC 1891

QY 11157 ATGAGCCCTGGATACACTGGCTTCTGAGCTGCACTTCT 11198

DB 1892 GTGCAGATTAGATCATTAGTTTAAACAGCTGCATCATAT 1933

RESULT 10

US-08-425-876-23

Sequence 23, Application US/08425876

Patent No. 6117422

GENERAL INFORMATION:

APPLICANT: LADNER, MARTHA B.

APPLICANT: NOBLE, JANELLE A.

APPLICANT: MARTIN, GEORGE A.

APPLICANT: KAWASAKI, ERNEST S.

APPLICANT: COYNE, MAZIE YEE

APPLICANT: HALENECK, ROBERT F.

APPLICANT: KOTHS, KIRSTON E.

TITLE OF INVENTION: NEW FORMS OF COLONY STIMULATING FACTOR-1

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHIRON CORPORATION

STREET: Intellectual Property - R440, P.O. Box 8097

CITY: Emeryville

STATE: California

COUNTRY: U.S.A.

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,876

FILING DATE: 21-APR-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McGarrigle Jr., Philip L.

REGISTRATION NUMBER: 31,395

REFERENCE/DOCKET NUMBER: 0681,011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 601-2718

TELEFAX: (510) 655-3542

TELEX: n/a

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1987 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 03:57:23 ; Search time 22135 Seconds

(without alignments)
15705.128 Million cell updates/sec

Title: US-09-922-549B-67

Perfect score: 11945

Sequence: 1 tgcgcctcttgatcattc.....tgcgcctgctgccttaggg 11945

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_prc:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_lnu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|--------------------|
| 1 | 11945 | 100.0 | 11945 | 5 | AF405540 | AF405540 Gallus ga |
| 2 | 11945 | 100.0 | 12728 | 5 | AF405538 | AF405538 Synthetic |
| 3 | 7974.4 | 66.8 | 18343 | 5 | AF410481 | AF410481 Gallus ga |
| 4 | 1374.4 | 11.5 | 1668 | 5 | GGBIH1 | X98408 G.gallus ly |
| 5 | 1374.4 | 11.5 | 4672 | 12 | CVE277960 | AJ277960 Cloning v |
| 6 | 644.2 | 5.4 | 660 | 5 | CICRIVAS | X52989 Chicken lys |
| 7 | 515.6 | 4.3 | 720 | 5 | CHRS25F | X06868 Gallus gall |
| 8 | 432.8 | 3.6 | 449 | 5 | GGDPAICP | X84223 G.gallus ly |
| 9 | 421 | 3.5 | 441 | 5 | GGIXSHE | X12509 Chicken lys |
| 10 | 390.4 | 3.3 | 2232 | 5 | CHRPOLIG | L22152 Gallus gall |
| 11 | 390.4 | 3.3 | 4558 | 5 | GGU88211 | U88211 Gallus gall |
| 12 | 382.4 | 3.2 | 1127 | 5 | CHRPOLID | L22149 Gallus gall |
| 13 | 377.6 | 3.2 | 1651 | 5 | CHRPOLIC | L22148 Gallus gall |
| 14 | 375.2 | 3.1 | 971 | 5 | CHKCR1A | M28069 Gallus gall |
| 15 | 375.2 | 3.1 | 3297 | 5 | GGVITCR1 | Y00324 Chicken vit |
| 16 | 366.2 | 3.1 | 98793 | 5 | AC091725 | AC091725 Gallus ga |
| 17 | 366.2 | 3.1 | 109569 | 5 | AC084760 | AC084760 Gallus ga |
| 18 | 352 | 2.9 | 117359 | 2 | AC123536 | AC123536 Gallus ga |
| 19 | 329.4 | 2.8 | 95651 | 2 | AC130796 | AC130796 Gallus ga |
| 20 | 320.8 | 2.7 | 84817 | 2 | AC126918 | AC126918 Gallus ga |
| 21 | 317 | 2.7 | 1623 | 5 | CHRPOLIB | L22147 Gallus gall |
| 22 | 315 | 2.6 | 347 | 12 | SYNLS2PR | M12532 Chicken lys |
| 23 | 313.6 | 2.6 | 31793 | 5 | GGU83833 | U83833 Gallus gall |
| 24 | 313.4 | 2.6 | 46304 | 5 | AF082667 | AF082667 Gallus ga |
| 25 | 309.6 | 2.6 | 123880 | 2 | AC117928 | AC117928 Gallus ga |
| 26 | 306.2 | 2.6 | 1238 | 5 | CHRPOLIA | L22146 Gallus gall |
| 27 | 295.8 | 2.5 | 167693 | 2 | AC120501 | AC120501 Gallus ga |
| 28 | 295 | 2.5 | 87496 | 2 | AC125497 | AC125497 Gallus ga |
| 29 | 294.8 | 2.5 | 123860 | 2 | AC117928 | AC117928 Gallus ga |
| 30 | 294.4 | 2.5 | 46022 | 5 | GGA289777 | AJ289777 Gallus ga |
| 31 | 294.4 | 2.5 | 46022 | 5 | GGA289778 | AJ289778 Gallus ga |
| 32 | 294.4 | 2.5 | 46032 | 5 | GGA289779 | AJ289779 Gallus ga |
| 33 | 293 | 2.5 | 214556 | 2 | AC120197 | AC120197 Gallus ga |
| 34 | 286.2 | 2.4 | 303 | 5 | GGLEY3 | X05463 Chicken lys |
| 35 | 285 | 2.4 | 136822 | 5 | AC091091 | AC091091 Gallus ga |
| 36 | 282 | 2.4 | 670 | 11 | AL592942 | AL592942 Meleagris |
| 37 | 281.2 | 2.4 | 113231 | 5 | AC094012 | AC094012 Gallus ga |
| 38 | 270.2 | 2.3 | 118914 | 5 | AC094011 | AC094011 Gallus ga |
| 39 | 268.8 | 2.3 | 27630 | 5 | AB042324 | AB042324 Gallus ga |
| 40 | 268.2 | 2.2 | 691 | 11 | AL592943 | AL592943 Meleagris |
| 41 | 258.6 | 2.2 | 170968 | 2 | AC118982 | AC118982 Gallus ga |
| 42 | 257.2 | 2.2 | 109569 | 5 | AC084760 | AC084760 Gallus ga |
| 43 | 253.4 | 2.1 | 6087 | 5 | AB052935 | AB052935 Gallus ga |
| 44 | 246.8 | 2.1 | 136116 | 2 | AC110874 | AC110874 Gallus ga |
| 45 | 243.4 | 2.0 | 31793 | 5 | GGU83833 | U83833 Gallus gall |

ALIGNMENTS

RESULT 1
AF405540 11945 bp DNA linear VRT 06-SEP-2001
LOCUS Gallus gallus lysozyme precursor, gene, partial cds.
DEFINITION AF405540
ACCESSION AF405540
VERSION AF405540.1 GI:15487991
KEYWORDS
SOURCE Gallus gallus.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus
REFERENCE 1 (bases 1 to 11945)
RAPP,J.C.
TITLE Chicken Lysozyme Promoter

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 11945)
AUTHORS Rapp, J.C.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2001) Avigenics Inc., 425 River Road, Athens,
Georgia 30602-2771, USA
FEATURES
Location/Qualifiers
Source 1. 11945
/organism="Gallus gallus"
/db_xref="taxon:9031"
misc_feature 1. 11891
/note="contains promoter and 5' UTR"
mRNA <11892..>11945
/product="lysozyme precursor"
11892..>11945
CDS /codon_start=1
/product="lysozyme precursor"
/protein_id="AA01041.1"
/db_xref="GI:15487992"
/translation="MRSLLILVLCFLPLALG"
sig_peptide 11892..11939
BASE COUNT 3123 a 2591 c 2630 g 3601 t
ORIGIN

Query Match 100.0%; Score 11945; DB 5; Length 11945;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 11945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCGCTCTCTTGAATTAATCACTGTTGATTTTCATCTCTTCCGCGATGAAGATA 60
DB 1 TGCCGCTCTCTTGAATTAATCACTGTTGATTTTCATCTCTTCCGCGATGAAGATA 60
QY 61 TAACAGCTGTATAACAGCTGTGAGGAATCTTGATTTTCTCTGATCAGTGTTTT 120
DB 61 TAACAGCTGTATAACAGCTGTGAGGAATCTTGATTTTCTCTGATCAGTGTTTT 120
QY 121 ATAAATATGTTGATATTTGATTAAGGCTGTGTCTTGTCTTGAGAGCAAAAGCCA 180
DB 121 ATAAATATGTTGATATTTGATTAAGGCTGTGTCTTGTCTTGAGAGCAAAAGCCA 180
QY 181 CAGCAGTGTTGGTGGGTGGTGGCAGCTCAGTACAGAGAGAGGTTTTTGGCTGTT 240
DB 181 CAGCAGTGTTGGTGGGTGGTGGCAGCTCAGTACAGAGAGAGGTTTTTGGCTGTT 240
QY 241 TTTTCTTTTTTTTTTTTAAAGAGTGTCTTTTCTTCTAGTAATTTTCTACTGA 300
DB 241 TTTTCTTTTTTTTTTTTAAAGAGTGTCTTTTCTTCTAGTAATTTTCTACTGA 300
QY 301 CTGTATGTTTGGACAGGTCAGAAACATTTCTCAAAAGAGAAAGCTTTGGAACTGTAC 360
DB 301 CTGTATGTTTGGACAGGTCAGAAACATTTCTCAAAAGAGAAAGCTTTGGAACTGTAC 360
QY 361 AGCCCTTTTCTTCAATCCCTTTTGTCTTGTGCAATGCTTGGTTCGATGTCAT 420
DB 361 AGCCCTTTTCTTCAATCCCTTTTGTCTTGTGCAATGCTTGGTTCGATGTCAT 420
QY 421 TATGGAAGAAAGCTTGAGCAACTGAGGTTTTTATTTAAGTGTGCTTGAAGCTTGA 480
DB 421 TATGGAAGAAAGCTTGAGCAACTGAGGTTTTTATTTAAGTGTGCTTGAAGCTTGA 480
QY 481 TAGCGTGTATACAGATACCTTATTAAGTTAGGCCAGCTTATGCTTATTTTTC 540
DB 481 TAGCGTGTATACAGATACCTTATTAAGTTAGGCCAGCTTATGCTTATTTTTC 540
QY 541 CCTTTGAGTAGTAGCTTCTCTGTTTCTTCTTGAAGCTGTTGAGCTTATGATTT 600
DB 541 CCTTTGAGTAGTAGCTTCTCTGTTTCTTCTTGAAGCTGTTGAGCTTATGATTT 600
QY 601 TTCTAATGAGATTTTAACTGATGATCTAGTTCATACCCAAATGCTTGAATGTTT 660
DB 601 TTCTAATGAGATTTTAACTGATGATCTAGTTCATACCCAAATGCTTGAATGTTT 660
QY 661 CCTAGTAAACATGTTGATTAACCTCGGATTTACATGTTGATATACTGTCATCTGT 720

DB 661 CCTAGTAAACATGTTGATTAACCTCGGATTTACATGTTGATATACTGTCATCTGT 720
QY 721 CTAGTAAATATATATGAGCTTTTATGAATFACGTATTCCTGATTTTATATAC 780
DB 721 CTAGTAAATATATATGAGCTTTTATGAATFACGTATTCCTGATTTTATATAC 780
QY 781 TCTATGCTGTGTGTACAGGTCAACAGACATTCCTATTTTATATATAGAAATTT 840
DB 781 TCTATGCTGTGTGTACAGGTCAACAGACATTCCTATTTTATATATAGAAATTT 840
QY 841 ATATGAGTCTGTCTGTTGGTCTTGTGTGTGAAGATATACAGCTTAAATTTCTAGAGCG 900
DB 841 ATATGAGTCTGTCTGTTGGTCTTGTGTGTGAAGATATACAGCTTAAATTTCTAGAGCG 900
QY 901 ATGCTCAGTAAAGCGGTTGTACATGGGTTCAAATGTAAACGGGACGTTGGCTGCT 960
DB 901 ATGCTCAGTAAAGCGGTTGTACATGGGTTCAAATGTAAACGGGACGTTGGCTGCT 960
QY 961 GCCTTCCGAGATCCAGACACTAAACCTCTCTGCACTGAGGTATTAATCCCTCAGAT 1020
DB 961 GCCTTCCGAGATCCAGACACTAAACCTCTCTGCACTGAGGTATTAATCCCTCAGAT 1020
QY 1021 CCCAGGGAAGTGCAGATCCAGCTGATTTCTTAAAGAAATGATTAATCTTAAATA 1080
DB 1021 CCCAGGGAAGTGCAGATCCAGCTGATTTCTTAAAGAAATGATTAATCTTAAATA 1080
QY 1081 TTTTGGCATATGAGAGCAACCTGATGATTTGTTGGACTTAAATTTATTTGTAACGG 1140
DB 1081 TTTTGGCATATGAGAGCAACCTGATGATTTGTTGGACTTAAATTTATTTGTAACGG 1140
QY 1141 AGTGCATAGGTTTTTAAACAGCTTCAGCATCTAAGCATGACAGACGTTTATGAGAG 1200
DB 1141 AGTGCATAGGTTTTTAAACAGCTTCAGCATCTAAGCATGACAGACGTTTATGAGAG 1200
QY 1201 TGATGCTGTGATGCTGTGTGAGCTGTTACGGCACTGCTTGCAAGTTCAGAT 1260
DB 1201 TGATGCTGTGATGCTGTGTGAGCTGTTACGGCACTGCTTGCAAGTTCAGAT 1260
QY 1261 AGGGGTGGGCTCTTGTGTCTGTTTCCACAGCTGACACAGCAACCTCCGGAACA 1320
DB 1261 AGGGGTGGGCTCTTGTGTCTGTTTCCACAGCTGACACAGCAACCTCCGGAACA 1320
QY 1321 CATCTACCTGTGGGTACTTTTCAAACCATCTTATGAGTATGATGATGATGATA 1380
DB 1321 CATCTACCTGTGGGTACTTTTCAAACCATCTTATGAGTATGATGATGATGATA 1380
QY 1381 ACAGAGAACTTCTCAGTTGGATATTTCTCATGGGATGCTTTTCCCATGTTGGGCAAA 1440
DB 1381 ACAGAGAACTTCTCAGTTGGATATTTCTCATGGGATGCTTTTCCCATGTTGGGCAAA 1440
QY 1441 GTATGATTAAGCATCTTATTTGTAATTTATGACATTTGTTAGTTCCTGATCTTCTAT 1500
DB 1441 GTATGATTAAGCATCTTATTTGTAATTTATGACATTTGTTAGTTCCTGATCTTCTAT 1500
QY 1501 AGCACCACTTATGAGAGAGGTGAGGCTGTGCTGCTGCTGCTTCAATCTT 1560
DB 1501 AGCACCACTTATGAGAGAGGTGAGGCTGTGCTGCTGCTGCTTCAATCTT 1560
QY 1561 TTAAGCTTCTTGAATATACATGATGATGTAAGTCTTGAAGATATGTAACACATA 1620
DB 1561 TTAAGCTTCTTGAATATACATGATGATGTAAGTCTTGAAGATATGTAACACATA 1620
QY 1621 CTATACCTTTGATCCCAATGAATTCAGAGATTTCAATTTGAAGAAATTCGCTATTCAT 1680
DB 1621 CTATACCTTTGATCCCAATGAATTCAGAGATTTCAATTTGAAGAAATTCGCTATTCAT 1680
QY 1681 ACCATGATATGATTTTAAACCCAGTGTGACACTTGGAAATATATCAAGTAATAG 1740
DB 1681 ACCATGATATGATTTTAAACCCAGTGTGACACTTGGAAATATATCAAGTAATAG 1740
QY 1741 ACTTTGGCTACCCCTTGTGTACTGATTTTGTAAATAGAAATATTTTAAACGTGCA 1800

Db 1741 ACTTGGCCCTGACCCCTCTGTCTACTGATTTTGTATAGAAAATTTTAACTGTGCA 1800
Qy 1801 TATGATTTATACATATATGAAGAGACATTTGCTGATCTTCAATATGTAAGAAATGAGA 1860
Db 1801 TATGATTTATACATATATGAAGAGACATTTGCTGATCTTCAATATGTAAGAAATGAGA 1860
Qy 1861 GTGGCTGTCTTTTAAATACAAAGTATGCAAAATGATGAGGTCTCTTAAAAAAA 1920
Db 1861 GTGGCTGTCTTTTAAATACAAAGTATGCAAAATGATGAGGTCTCTTAAAAAAA 1920
Qy 1921 AAAAAAAGATATATTAAGAGACAGGTGTTTACAAGTGAATACATTTCTATTGG 1980
Db 1921 AAAAAAAGATATATTAAGAGACAGGTGTTTACAAGTGAATACATTTCTATTGG 1980
Qy 1981 TAAACAGTTACATTTTATGAAGATTACAGGCGTGTGACTTTCTAAACATTAAGGCTGT 2040
Db 1981 TAAACAGTTACATTTTATGAAGATTACAGGCGTGTGACTTTCTAAACATTAAGGCTGT 2040
Qy 2041 ATTGCTCTCTGTACCATTTGCTCATTTCCCAATTTGACAGAGATGTCTGGTAA 2100
Db 2041 ATTGCTCTCTGTACCATTTGCTCATTTCCCAATTTGACAGAGATGTCTGGTAA 2100
Qy 2101 ACTATTCAGAAATGCTTTGAATACAGCATGGAGCTTGTCTGATTTGAGATCAGAG 2160
Db 2101 ACTATTCAGAAATGCTTTGAATACAGCATGGAGCTTGTCTGATTTGAGATCAGAG 2160
Qy 2161 TTGCACTGCAAAATGTCAGSAAATGATGTCTGCAAGATGCCAATCCAAAGGATTTT 2220
Db 2161 TTGCACTGCAAAATGTCAGSAAATGATGTCTGCAAGATGCCAATCCAAAGGATTTT 2220
Qy 2221 ATATGTATATATAGTAAGCAGTTTCTGATTCACAGAGCCCAAGAGTCTGCTGATCTT 2280
Db 2221 ATATGTATATATAGTAAGCAGTTTCTGATTCACAGAGCCCAAGAGTCTGCTGATCTT 2280
Qy 2281 GTGTGCGCGAGACCTGATTTCTTAACAAGGTAGATGATCTCTAGCAACTGGGATTT 2340
Db 2281 GTGTGCGCGAGACCTGATTTCTTAACAAGGTAGATGATCTCTAGCAACTGGGATTT 2340
Qy 2341 TTAATACATTTTACAGCAAGACTTATTAATCTCTACCTTGAAGATCGTTTATCAT 2400
Db 2341 TTAATACATTTTACAGCAAGACTTATTAATCTCTACCTTGAAGATCGTTTATCAT 2400
Qy 2401 TTTTATAGATGTTATATCTGAAATACGTAACCTTTTACCTTCAATGAGGTTCTTTTTC 2460
Db 2401 TTTTATAGATGTTATATCTGAAATACGTAACCTTTTACCTTCAATGAGGTTCTTTTTC 2460
Qy 2461 AGCCTTGAAGAGACTGTTAAGCAATTTGCTGTCCAATTTTGTGTCTTAACTGCA 2520
Db 2461 AGCCTTGAAGAGACTGTTAAGCAATTTGCTGTCCAATTTTGTGTCTTAACTGCA 2520
Qy 2521 ATAGTAGTTTACCTGTATTTGAGAAATTAAGACCAATTTTATTAATAAATACCTTTTG 2580
Db 2521 ATAGTAGTTTACCTGTATTTGAGAAATTAAGACCAATTTTATTAATAAATACCTTTTG 2580
Qy 2581 TCTGTCTCATTTTGCATTTGCTGATATCTTTGACAGTCCCATTTATGTCAGTCTGTCAG 2640
Db 2581 TCTGTCTCATTTTGCATTTGCTGATATCTTTGACAGTCCCATTTATGTCAGTCTGTCAG 2640
Qy 2641 ATATTTCAGACATCAAAACTTAACTGAGCTCAGTGTACAGCTGGGTTTATGCT 2700
Db 2641 ATATTTCAGACATCAAAACTTAACTGAGCTCAGTGTACAGCTGGGTTTATGCT 2700
Qy 2701 GTTATTTATTTTGAAGTGAAGATGATGTGTCTCATCTGCTCATCAACCTTCATGTC 2760
Db 2701 GTTATTTATTTTGAAGTGAAGATGATGTGTCTCATCTGCTCATCAACCTTCATGTC 2760
Qy 2761 AGAGTGAAGCTAGTGAAGATGATGATTTATTTGATTTTAAAGTCAACTTTT 2820
Db 2761 AGAGTGAAGCTAGTGAAGATGATGATTTATTTGATTTTAAAGTCAACTTTT 2820
Qy 2821 TATCAGATTTTATTTTCAATTTGGAATATATTTGTTTCTAGACGTCATAGCTTGAATC 2880
Db 2821 TATCAGATTTTATTTTCAATTTGGAATATATTTGTTTCTAGACGTCATAGCTTGAATC 2880

Qy 2881 TGAATGCACTGTGATTTGTCATGTAAGAACACAGCACTTCTCATTTACTTAACTTCAT 2940
Db 2881 TGAATGCACTGTGATTTGTCATGTAAGAACACAGCACTTCTCATTTACTTAACTTCAT 2940
Qy 2941 TTTGGAATGCAAGGATTAAGCAAGGGCACAGTCTCATGTAATATAGACAGTGCCTCAG 3000
Db 2941 TTTGGAATGCAAGGATTAAGCAAGGGCACAGTCTCATGTAATATAGACAGTGCCTCAG 3000
Qy 3001 GAGAAAGTGAACCTGGATTTCTTTGGCTAGTGTCTTAAATCTGTAGTGAAGAAATACA 3060
Db 3001 GAGAAAGTGAACCTGGATTTCTTTGGCTAGTGTCTTAAATCTGTAGTGAAGAAATACA 3060
Qy 3061 CCCGATTCCTTGAAGGGCTCCAGCTTTATGCTTCCAAATTTGAAGGTGGCGGCACTT 3120
Db 3061 CCCGATTCCTTGAAGGGCTCCAGCTTTATGCTTCCAAATTTGAAGGTGGCGGCACTT 3120
Qy 3121 GGCACACTGGTTATTTACTGCAATTAATGTCACAGTTTCCGACGTAACTGGCTTCTCACATA 3180
Db 3121 GGCACACTGGTTATTTACTGCAATTAATGTCACAGTTTCCGACGTAACTGGCTTCTCACATA 3180
Qy 3181 TTGACATGCACTATATAGCTTGGCTTCAGAGCCAGGTGAAGTTGGGATGGTGAAGGA 3240
Db 3181 TTGACATGCACTATATAGCTTGGCTTCAGAGCCAGGTGAAGTTGGGATGGTGAAGGA 3240
Qy 3241 GTGCTGGGCTGTGGCTGGGGGAGACTGTGGGAGCTCCAAAGCTGAGCTGGGGGTGGCAGCA 3300
Db 3241 GTGCTGGGCTGTGGCTGGGGGAGACTGTGGGAGCTCCAAAGCTGAGCTGGGGGTGGCAGCA 3300
Qy 3301 CAGGAAAAAGTGTGGGTAATATTTTAAATGATCTGTGTGCAAAAGCTCATCTCAAT 3360
Db 3301 CAGGAAAAAGTGTGGGTAATATTTTAAATGATCTGTGTGCAAAAGCTCATCTCAAT 3360
Qy 3361 AGTAGGGGTGTACTCTGCAAGATTAACAGTGTGGTCACTAATATATGTAATATTC 3420
Db 3361 AGTAGGGGTGTACTCTGCAAGATTAACAGTGTGGTCACTAATATATGTAATATTC 3420
Qy 3421 ACAGTGAAGCAATTCAGAGGTATCATTAACGACACACAGATCATCAATGATATTGG 3480
Db 3421 ACAGTGAAGCAATTCAGAGGTATCATTAACGACACACAGATCATCAATGATATTGG 3480
Qy 3481 AAGCGTATCAGAGAGCGAGAGGTAAAGCACTTCTATATGTTTTCCCTCCAGATAA 3540
Db 3481 AAGCGTATCAGAGAGCGAGAGGTAAAGCACTTCTATATGTTTTCCCTCCAGATAA 3540
Qy 3541 GCAGTCTGGGAAAGTAGCACCCCTTGAGCAGAGACAAAGAAATTAATTCAGAGCATGTGC 3600
Db 3541 GCAGTCTGGGAAAGTAGCACCCCTTGAGCAGAGACAAAGAAATTAATTCAGAGCATGTGC 3600
Qy 3601 TAGAGAACTTTCTTGCTGAATTTCTACTTGAAGAGCTTTGATGCTGTGCTGTGCTGC 3660
Db 3601 TAGAGAACTTTCTTGCTGAATTTCTACTTGAAGAGCTTTGATGCTGTGCTGTGCTGC 3660
Qy 3661 TTTCTGACGACCTGCAAGGCCAGAGCCCTGTGGTGAAGTGAAGGAAAGATTTCTGCTCA 3720
Db 3661 TTTCTGACGACCTGCAAGGCCAGAGCCCTGTGGTGAAGTGAAGGAAAGATTTCTGCTCA 3720
Qy 3721 GTTCAAGCTTCAGCAGGTATTTGCTTTTCTTCCACACACTGTGCACAGAGTGG 3780
Db 3721 GTTCAAGCTTCAGCAGGTATTTGCTTTTCTTCCACACACTGTGCACAGAGTGG 3780
Qy 3781 AACTGATGTGCAAGGCTCTGTCACACTACCTTTGTGAGGAGCAGACGCTCTCGAANA 3840
Db 3781 AACTGATGTGCAAGGCTCTGTCACACTACCTTTGTGAGGAGCAGACGCTCTCGAANA 3840
Qy 3841 AGAGAGCTAACTCTATGTCATATGTCGAAGGTAAGTGGTTTTTAAAAAAGAAACACAA 3900
Db 3841 AGAGAGCTAACTCTATGTCATATGTCGAAGGTAAGTGGTTTTTAAAAAAGAAACACAA 3900
Qy 3901 AGGCAAAACCGGCTGCCCATGAGAAAGCAGTGTGAATATGTTAGAAAAGTGCAG 3960
Db 3901 AGGCAAAACCGGCTGCCCATGAGAAAGCAGTGTGAATATGTTAGAAAAGTGCAG 3960

OY 3961 AAGCCCCAGGACGTGTGACAGGCCCTCTGCCACCTAGAGGCGGAACAAGCTTCCCT 4020
| | | | |
Db 3961 AAGCCCCAGGACGTGTGACAGGCCCTCTGCCACCTAGAGGCGGAACAAGCTTCCCT 4020
OY 4021 GCGTAGGGCTGTGGCCGGGAGAGTGCGTTCTTGGTGGGCTTTTGGTGGCGTTGGCT 4080
| | | | |
Db 4021 GCGTAGGGCTGTGGCCGGGAGAGTGCGTTCTTGGTGGGCTTTTGGTGGCGTTGGCT 4080
OY 4081 TTGAGATTTAGACACAAGGAGAGCTGAAAGAGAGTGTGGGCACTATTTTGGTTGTA 4140
| | | | |
Db 4081 TTGAGATTTAGACACAAGGAGAGCTGAAAGAGAGTGTGGGCACTATTTTGGTTGTA 4140
OY 4141 AGCGTGTACTTCAAAATATATTTTGTGAGGAGGTGTAAGCAATTTGCCAATTTAAATA 4200
| | | | |
Db 4141 AGCGTGTACTTCAAAATATATTTTGTGAGGAGGTGTAAGCAATTTGCCAATTTAAATA 4200
OY 4201 AAGTTGCAAGAGATTGAAGGCTGAGTAGTAGAGGTAACACGTTTATATGATCTTCT 4260
| | | | |
Db 4201 AAGTTGCAAGAGATTGAAGGCTGAGTAGTAGAGGTAACACGTTTATATGATCTTCT 4260
OY 4261 GAAACTACTGCTTCAAAACACTTGTGTGAGTGTGAGACCTTGGATAGTAGAGTCTCT 4320
| | | | |
Db 4261 GAAACTACTGCTTCAAAACACTTGTGTGAGTGTGAGACCTTGGATAGTAGAGTCTCT 4320
OY 4321 GTTACATGTCTGATGACACTTGTCTTGTCCATCCACATCCATGCACTTCCATCCA 4380
| | | | |
Db 4321 GTTACATGTCTGATGACACTTGTCTTGTCCATCCACATCCATGCACTTCCATCCA 4380
OY 4381 CGCATTTGTCACTTATCCCATATCTGTCATCTGACATCTGTCCTTGTGTCACCTGG 4440
| | | | |
Db 4381 CGCATTTGTCACTTATCCCATATCTGTCATCTGACATCTGTCCTTGTGTCACCTGG 4440
OY 4441 TCAGAGAAACAGATGTGATATATCCACGCCGCCCAAGTTGAGAAGATGCGAGTTGCT 4500
| | | | |
Db 4441 TCAGAGAAACAGATGTGATATATCCACGCCGCCCAAGTTGAGAAGATGCGAGTTGCT 4500
OY 4501 TCTTCCCTTTTCTGCTAGTAGTAGATTTTCTCTGGCTTTGACACCTACAGCAATAG 4560
| | | | |
Db 4501 TCTTCCCTTTTCTGCTAGTAGTAGATTTTCTCTGGCTTTGACACCTACAGCAATAG 4560
OY 4561 TCTTCTGCTTACATTTCTGGGCACTTATTTCAAAATATCTTGGAGTGGCTGCTCTCAG 4620
| | | | |
Db 4561 TCTTCTGCTTACATTTCTGGGCACTTATTTCAAAATATCTTGGAGTGGCTGCTCTCAG 4620
OY 4621 TTTGTGTCTTCTACTCTTAGAGTGAATGCTTGAAGTGAAGAGAGAGAGAT 4680
| | | | |
Db 4621 TTTGTGTCTTCTACTCTTAGAGTGAATGCTTGAAGTGAAGAGAGAGAGAT 4680
OY 4681 GTTGGCGGAGTTCTGATGAACACACCTGTAATATATGSCCAAGGTGGGTGGTTTC 4740
| | | | |
Db 4681 GTTGGCGGAGTTCTGATGAACACACCTGTAATATATGSCCAAGGTGGGTGGTTTC 4740
OY 4741 TCTGAGGAACGGGAGGTTTGTCTGTGAAGCAAGAGACTCTGCGAGTTGCAATATT 4800
| | | | |
Db 4741 TCTGAGGAACGGGAGGTTTGTCTGTGAAGCAAGAGACTCTGCGAGTTGCAATATT 4800
OY 4801 TTGCAACTGATGTGGAAGTGTGCTTAAAGCAAGTTCCCTAGTTCCTCTACTTCTT 4860
| | | | |
Db 4801 TTGCAACTGATGTGGAAGTGTGCTTAAAGCAAGTTCCCTAGTTCCTCTACTTCTT 4860
OY 4861 TTCTCTTCTGGCAGTCAATTTATTTCTGACAGACAACAGCACCACCTGAGGCTTA 4920
| | | | |
Db 4861 TTCTCTTCTGGCAGTCAATTTATTTCTGACAGACAACAGCACCACCTGAGGCTTA 4920
OY 4921 GAAAGTATGTGGCTCTGCGTGGGTGTGTACAGCTTGCCTGGTGAAGGGAATTAATA 4980
| | | | |
Db 4921 GAAAGTATGTGGCTCTGCGTGGGTGTGTACAGCTTGCCTGGTGAAGGGAATTAATA 4980
OY 4981 CGGCGACCATTTCAATCCCAAGAGATCTCTCATTTATGATGAATGAAGTGAAGTGGG 5040
| | | | |
Db 4981 CGGCGACCATTTCAATCCCAAGAGATCTCTCATTTATGATGAATGAAGTGAAGTGGG 5040
OY 5041 CTCACACTCAAAACATTAATTGAGTAGAAGATGAATTAATAACGCAATTTCGCAATTCC 5100
| | | | |

Db 5041 CTCACACTCAAAACATTAATTGAGTAGAAGTGAATTAATAACGCAATTTCGCAATTCC 5100
| | | | |
OY 5101 TAAGCATTTAGTGTGAGCTGTGACAGATGTAGAGTGGGCACTCCACTTCTCAAAAG 5160
| | | | |
Db 5101 TAAGCATTTAGTGTGAGCTGTGACAGATGTAGAGTGGGCACTCCACTTCTCAAAAG 5160
OY 5161 CACTGATGAGAGTAGTAAATAATGAGACGATTCAGAACCAACAGGAGTGTGGCG 5220
| | | | |
Db 5161 CACTGATGAGAGTAGTAAATAATGAGACGATTCAGAACCAACAGGAGTGTGGCG 5220
OY 5221 AAGAACTGATGGAATATATGATGAATTTGTGTGTGACATTTTTTTTAAATACATPAA 5280
| | | | |
Db 5221 AAGAACTGATGGAATATATGATGAATTTGTGTGTGACATTTTTTTTAAATACATPAA 5280
OY 5281 CTACTTCAATGAGTGTGAGAGTGTGATGTTTATATAGAGCATTAACCCAGGTGAG 5340
| | | | |
Db 5281 CTACTTCAATGAGTGTGAGAGTGTGATGTTTATATAGAGCATTAACCCAGGTGAG 5340
OY 5341 CGAGTACCATTTTCTTACAAAGAAACGATTCGAGCTGTGGTAAATAGATTCTC 5400
| | | | |
Db 5341 CGAGTACCATTTTCTTACAAAGAAACGATTCGAGCTGTGGTAAATAGATTCTC 5400
OY 5401 CATAGCGCTGAAGTCCCTGCTGCTGCCATCTCAGCTGAGTGAAGTGCATTT 5460
| | | | |
Db 5401 CATAGCGCTGAAGTCCCTGCTGCTGCCATCTCAGCTGAGTGAAGTGCATTT 5460
OY 5461 CTTGGGGTTCCTGACAGAGTAAATGGAATACTTCAAAAATCTTCTTTTCC 5520
| | | | |
Db 5461 CTTGGGGTTCCTGACAGAGTAAATGGAATACTTCAAAAATCTTCTTTTCC 5520
OY 5521 TGTCAATGTGGATCCCTACTGTGCTCTGCTTGTACGTTACCCCTGACTGTTCCATT 5580
| | | | |
Db 5521 TGTCAATGTGGATCCCTACTGTGCTCTGCTTGTACGTTACCCCTGACTGTTCCATT 5580
OY 5581 CAGCGTTTGAAGAGAAAAAATTTGAAATTAACATGTCTAGCTTTCACCTCT 5640
| | | | |
Db 5581 CAGCGTTTGAAGAGAAAAAATTTGAAATTAACATGTCTAGCTTTCACCTCT 5640
OY 5641 CCAGCATTTTGTGTTTAAATATGCAATTAACCTGCTTGAATGGAATGAGAGGGGT 5700
| | | | |
Db 5641 CCAGCATTTTGTGTTTAAATATGCAATTAACCTGCTTGAATGGAATGAGAGGGGT 5700
OY 5701 TGGGTGTATTACCGAGAACAAAGAGGCTTATATTAACCTGCTTATTTAGAGA 5760
| | | | |
Db 5701 TGGGTGTATTACCGAGAACAAAGAGGCTTATATTAACCTGCTTATTTAGAGA 5760
OY 5761 ACTGGCAAGCTGTCAAAAACAAAAGGCTTACCAACCAATTAAGTGAATGCGGCTATA 5820
| | | | |
Db 5761 ACTGGCAAGCTGTCAAAAACAAAAGGCTTACCAACCAATTAAGTGAATGCGGCTATA 5820
OY 5821 GCCAGCAGGGCAGCAGAGGAGTGTGCACTGTGCACTATGCAAGGCTGCTGTGTG 5880
| | | | |
Db 5821 GCCAGCAGGGCAGCAGAGGAGTGTGCACTGTGCACTATGCAAGGCTGCTGTGTG 5880
OY 5881 ACTCTGAGAGCAACTGCTTGTGAATGACACACTTGTGCAATTTCTTGTTCAGAA 5940
| | | | |
Db 5881 ACTCTGAGAGCAACTGCTTGTGAATGACACACTTGTGCAATTTCTTGTTCAGAA 5940
OY 5941 TGGGTAGAGCGTGTGCTGGGAGAGTTTCTGATTTAGGCGACCTGTTTCTCTCTC 6000
| | | | |
Db 5941 TGGGTAGAGCGTGTGCTGGGAGAGTTTCTGATTTAGGCGACCTGTTTCTCTCTC 6000
OY 6001 TCTTCATTTCTCTAAGCATGTCTCATGCTGTGTAATCCAGTCAAGTGAAGTTCACAA 6060
| | | | |
Db 6001 TCTTCATTTCTCTAAGCATGTCTCATGCTGTGTAATCCAGTCAAGTGAAGTTCACAA 6060
OY 6061 ATGAATCATCATCTGTAGAGTTCGTGTGTATCAAAATCTTGTGTGAGTCTATAAAT 6120
| | | | |
Db 6061 ATGAATCATCATCTGTAGAGTTCGTGTGTATCAAAATCTTGTGTGAGTCTATAAAT 6120
OY 6121 ATGGAAGCTTATTTATTTTGTCTTCCATATCAAGTCTTCTATGCAATTCACATCC 6180
| | | | |

Dh 6121 ATGGAAGCTTATTTATTTTTCGTTCTTCATATCATGCTCTTCTATGACATTCACATCC 6180
Qy 6181 ACCACAGCAATTAATAAGGTGAAGAGGCTGGTGGATGAAGAGGCTCTTACGTTTACG 6240
Dh 6181 ACCACAGCAATTAATAAGGTGAAGAGGCTGGTGGATGAAGAGGCTCTTACGTTTACG 6240
Qy 6241 TTCTTCTTCGAAGGCGACAGGAATAATGCTAGAGCTGTAGATATACAGCCCTGGGGTAGA 6300
Dh 6241 TTCTTCTTCGAAGGCGACAGGAATAATGCTAGAGCTGTAGATATACAGCCCTGGGGTAGA 6300
Qy 6301 AGTTCACTCTCTGCTGGAGACGCTAACCGCATCTTATACCCCTCTGAGACTCATCTT 6360
Dh 6301 AGTTCACTCTCTGCTGGAGACGCTAACCGCATCTTATACCCCTCTGAGACTCATCTT 6360
Qy 6361 AGGACCAATAAGGCTCATCTGGGGTTTTTGTCTGCTGCTCTCTGCTGGAAGGCTATCT 6420
Dh 6361 AGGACCAATAAGGCTCATCTGGGGTTTTTGTCTGCTGCTCTCTGCTGGAAGGCTATCT 6420
Qy 6421 CACTATTTTCACTGCTCCACGCTTACAAACCAAGATACAGCCGTGAATTTTTCGTAGGCC 6480
Dh 6421 CACTATTTTCACTGCTCCACGCTTACAAACCAAGATACAGCCGTGAATTTTTCGTAGGCC 6480
Qy 6481 ACATTACATAAATTTGACCTGTGTACCAATATGTTCTCTATATAGTTATTTCTTCCCA 6540
Dh 6481 ACATTACATAAATTTGACCTGTGTACCAATATGTTCTCTATATAGTTATTTCTTCCCA 6540
Qy 6541 CTGTGTTTAAACCCCTTAGGCAATTCAGAACACTAGAAATCATAGATGGTTGGATTTGA 6600
Dh 6541 CTGTGTTTAAACCCCTTAGGCAATTCAGAACACTAGAAATCATAGATGGTTGGATTTGA 6600
Qy 6601 AGGGGCTTTAAACATCATCCATTTTCAACCCCTGACACCTGCTTGGGCTTGGCCACCCTG 6660
Dh 6601 AGGGGCTTTAAACATCATCCATTTTCAACCCCTGACACCTGCTTGGGCTTGGCCACCCTG 6660
Qy 6661 CTGAGGCTGGCCAGAGGCTGCTGCAACCTGAGCCCTGAGACCTTCAGAGGATGGGGCACC 6720
Dh 6661 CTGAGGCTGGCCAGAGGCTGCTGCAACCTGAGCCCTGAGACCTTCAGAGGATGGGGCACC 6720
Qy 6721 ACAGCTTCTGCTGGGAGAGCTGTGCAACACTGACACTCTGCTGGGTAAAGATTTCTCT 6780
Dh 6721 ACAGCTTCTGCTGGGAGAGCTGTGCAACACTGACACTCTGCTGGGTAAAGATTTCTCT 6780
Qy 6781 TTAAACATCTAATCTAATCTCTCTCTTTTAAAGCCATCTCTTTTCCCGTTG 6840
Dh 6781 TTAAACATCTAATCTAATCTCTCTCTTTTAAAGCCATCTCTTTTCCCGTTG 6840
Qy 6841 CTATCTGTCCAGAAATGTGTATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6900
Dh 6841 CTATCTGTCCAGAAATGTGTATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6900
Qy 6901 GCGTGCAGTGAAGTCTGCCCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6960
Dh 6901 GCGTGCAGTGAAGTCTGCCCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6960
Qy 6961 CAGCCTGTCTCTGAGAGATCATCTTAAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7020
Dh 6961 CAGCCTGTCTCTGAGAGATCATCTTAAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7020
Qy 7021 CAGGCTTTCTGAGAGGCT 7080
Dh 7021 CAGGCTTTCTGAGAGGCT 7080
Qy 7081 GCAGAGCAGATGAGGAGAAATGCTTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7140
Dh 7081 GCAGAGCAGATGAGGAGAAATGCTTACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7140
Qy 7141 CCCAGGCTACTGTTGAGGCT 7200
Dh 7141 CCCAGGCTACTGTTGAGGCT 7200
Qy 7201 CATTCACAGAGACCGAGCT 7260
Dh 7201 CATTCACAGAGACCGAGCT 7260

Qy 7261 TCAGAGACTTCACATCTTTAGAGACAGCTGTGTACACCTACCTGCCCCATTTCTGTGAT 7320
Dh 7261 TCAGAGACTTCACATCTTTAGAGACAGCTGTGTACACCTACCTGCCCCATTTCTGTGAT 7320
Qy 7321 ATATACATTTTCAATGCTTCT 7380
Dh 7321 ATATACATTTTCAATGCTTCT 7380
Qy 7381 TACATGCAAGATTTCT 7440
Dh 7381 TACATGCAAGATTTCT 7440
Qy 7441 CAATTTGCTGCAATACCTTCCAGCTGCGGCGCCCTCCATTAATCTCTATTTGGGATCAG 7500
Dh 7441 CAATTTGCTGCAATACCTTCCAGCTGCGGCGCCCTCCATTAATCTCTATTTGGGATCAG 7500
Qy 7501 TTACCTTTTGGGGTAAAGCTTTTGTATGTCGAGAGACCTGCGGGCTTGTGATGCTCTCAG 7560
Dh 7501 TTACCTTTTGGGGTAAAGCTTTTGTATGTCGAGAGACCTGCGGGCTTGTGATGCTCTCAG 7560
Qy 7561 CTCTGCTCTGTTGACTGACACCAATTTTCTAGATCACCAGTGTCTCTGTAACAATTC 7620
Dh 7561 CTCTGCTCTGTTGACTGACACCAATTTTCTAGATCACCAGTGTCTCTGTAACAATTC 7620
Qy 7621 TTGTCTCTCATCTCTTCCAGCTGTATCTTGTGCAAAATACAGGCTATTTTGTGTTTTG 7680
Dh 7621 TTGTCTCTCATCTCTTCCAGCTGTATCTTGTGCAAAATACAGGCTATTTTGTGTTTTG 7680
Qy 7681 CTTCAGAGCCATTTAATTTCT 7740
Dh 7681 CTTCAGAGCCATTTAATTTCT 7740
Qy 7741 TTTTCAGAGCTTGTGCAAGAAACATCTAGCTGAAACCTTTCTGCAATTTCTTAC 7800
Dh 7741 TTTTCAGAGCTTGTGCAAGAAACATCTAGCTGAAACCTTTCTGCAATTTCTTAC 7800
Qy 7801 CAGTTCTCTCTGTTGAGTGAAGCATTAATTTACTAGAACTTCTGCACTGCAAGTTTAT 7860
Dh 7801 CAGTTCTCTCTGTTGAGTGAAGCATTAATTTACTAGAACTTCTGCACTGCAAGTTTAT 7860
Qy 7861 GCATTTTATTTACTTATTTATGACTTACTTGTGACATTAACAGACGACATATTTTG 7920
Dh 7861 GCATTTTATTTACTTATTTATGACTTACTTGTGACATTAACAGACGACATATTTTG 7920
Qy 7921 CTGGGATTTCCACAGTGTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7980
Dh 7921 CTGGGATTTCCACAGTGTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7980
Qy 7981 AACCTTGGCAATCTGCCCCAGCTGCCCCATCACAGAAAGAGATTTCTTTTATTTACTTC 8040
Dh 7981 AACCTTGGCAATCTGCCCCAGCTGCCCCATCACAGAAAGAGATTTCTTTTATTTACTTC 8040
Qy 8041 TCTTCAGCCATTAACAAATGTGAGAGGCCCAACAGAACTGTGGGCGAGGCTGCA 8100
Dh 8041 TCTTCAGCCATTAACAAATGTGAGAGGCCCAACAGAACTGTGGGCGAGGCTGCA 8100
Qy 8101 TCAAGGAGAGACAGCTGTAAGGTTGTGTACTCAATTAAGATTAAGAAATTAAGAGT 8160
Dh 8101 TCAAGGAGAGACAGCTGTAAGGTTGTGTACTCAATTAAGATTAAGAAATTAAGAGT 8160
Qy 8161 TGTGAGACAGTTTGTGCTGATTTATACAGGAGGCCCAACGAGAGGCTGTCTGCA 8220
Dh 8161 TGTGAGACAGTTTGTGCTGATTTATACAGGAGGCCCAACGAGAGGCTGTCTGCA 8220
Qy 8221 AGGCGACCTTGCACCTCTTGTGTTGTAAGATTAAGTAACTTCTTCTGTGAATTG 8280
Dh 8221 AGGCGACCTTGCACCTCTTGTGTTGTAAGATTAAGTAACTTCTTCTGTGAATTG 8280
Qy 8281 CGTGAGAGATCATGATGGCACTTCTGTGTTTACTATGTAAGATGTAAGATTAAGAGAG 8340
Dh 8281 CGTGAGAGATCATGATGGCACTTCTGTGTTTACTATGTAAGATGTAAGATTAAGAGAG 8340

| | | | |
|----|------|---|------|
| QY | 8341 | CAGAAAGTAACACTTGCTGCTGCTAGTGTGCTGCTATCCAGACAGGATGCGACTGCA | 8400 |
| Db | 8341 | CAGCAAGTAACCTTGCTGCTGCTAGTGTGCTGCTATCCAGACAGGATGCGACTGCA | 8400 |
| QY | 8401 | CACCAAGATAGAGGATGCTGCCAGCTGACGAGATGCTGGGACAGTAACAGTGGGCCATG | 8460 |
| Db | 8401 | CACCAAGATAGAGGATGCTGCCAGCTGACGAGATGCTGGGACAGTAACAGTGGGCCATG | 8460 |
| QY | 8461 | CTGCGTGTCAATPAGCATCACTCAGCCCTCACAGCCCTCAGAAAGATCATCCCAAG | 8520 |
| Db | 8461 | CTGCGCTCATATPAGCATCACTCAGCCCTCACAGCCCTCAGAAAGATCATCCCAAG | 8520 |
| QY | 8521 | TCAGGAAACTGCTCATCTCTTCCACATCATCAAACTTTGGGCTGACTATGCTGCCG | 8580 |
| Db | 8521 | TCAGGAAACTGCTCATCTCTTCCACATCATCAAACTTTGGGCTGACTATGCTGCCG | 8580 |
| QY | 8581 | GATGCTTAAATGTGTGCTACGTGACATCTTATTTTCTATGATTCAAGTCAGAACCTCG | 8640 |
| Db | 8581 | GATGCTTAAATGTGTGCTACGTGACATCTTATTTTCTATGATTCAAGTCAGAACCTCG | 8640 |
| QY | 8641 | GATCAGAGAGGAACACATATGTGGGAATGTACCTTCAGCTCCAAAGCCAGATCTTCTTCA | 8700 |
| Db | 8641 | GATCAGAGAGGAACACATATGTGGGAATGTACCTTCAGCTCCAAAGCCAGATCTTCTTCA | 8700 |
| QY | 8701 | ATGATCATCATGCTCTATTAGAAAGTGTGTGTGAATGTGAATTTGCCCTTTGTAT | 8760 |
| Db | 8701 | ATGATCATCATGCTCTATTAGAAAGTGTGTGTGAATGTGAATTTGCCCTTTGTAT | 8760 |
| QY | 8761 | TTTTTCTCTGCTGTACAGAAACATTTTGAATCCAGAGAAAAAGTAAGTGCCTCTT | 8820 |
| Db | 8761 | TTTTTCTCTGCTGTACAGAAACATTTTGAATCCAGAGAAAAAGTAAGTGCCTCTTCTT | 8820 |
| QY | 8821 | GGCATGGAGAGATTTGTCACTTGCAGAAAAATAAGATGCGATCCCAATGTTCATATATC | 8880 |
| Db | 8821 | GGCATGGAGAGATTTGTCACTTGCAGAAAAATAAGATGCGATCCCAATGTTCATATATC | 8880 |
| QY | 8881 | TCAGGCTCGAAGGAGATCAGAAACTGTGTATACAAATTTCAGGCTTCTGTGAATGCAGC | 8940 |
| Db | 8881 | TCAGGCTCGAAGGAGATCAGAAACTGTGTATACAAATTTCAGGCTTCTGTGAATGCAGC | 8940 |
| QY | 8941 | TTTTTGAACCTGTCTCTGGCCGAGCAGTACTGTAGTAAGAACCTCGGAAACAGAGAACAA | 9000 |
| Db | 8941 | TTTTTGAACCTGTCTCTGGCCGAGCAGTACTGTAGTAAGAACCTCGGAAACAGAGAACAA | 9000 |
| QY | 9001 | TCTCTTCAAGTGTACAGCAGAGAGAAACACCTTGGCCATCATGAAGATGAATTAACACTGC | 9060 |
| Db | 9001 | TCTCTTCAAGTGTACAGCAGAGAGAAACACCTTGGCCATCATGAAGATGAATTAACACTGC | 9060 |
| QY | 9061 | CGCTGAAGAAATCCACGCTCTGTTTATAGAGAGTGTCACACCTCCACACTGGAACACACA | 9120 |
| Db | 9061 | CGCTGAAGAAATCCACGCTCTGTTTATAGAGAGTGTCACACCTCCACACTGGAACACACA | 9120 |
| QY | 9121 | GTTTCATTTTATAGACTTCCAGAGAGATCTTCTTAAGCTTCTTAATTATGTGACA | 9180 |
| Db | 9121 | GTTTCATTTTATAGACTTCCAGAGAGATCTTCTTAAGCTTCTTAATTATGTGACA | 9180 |
| QY | 9181 | TCTCCAGTTGGCAGATGACTATGACTACTGACAGAGAAATGAGAACTAGCTGGGAATAT | 9240 |
| Db | 9181 | TCTCCAGTTGGCAGATGACTATGACTACTGACAGAGAAATGAGAACTAGCTGGGAATAT | 9240 |
| QY | 9241 | TTCCTGTTTACCAACATGAGATGCCCATTTTCTTACTGTAATTGGAATTAATATCT | 9300 |
| Db | 9241 | TTCCTGTTTACCAACATGAGATGCCCATTTTCTTACTGTAATTGGAATTAATATCT | 9300 |
| QY | 9301 | GAAATTCGAAGCAGGATAGGGAATATCTCAATTTCTTCCATGTTGTGTACAGCACAGT | 9360 |
| Db | 9301 | GAAATTCGAAGCAGGATAGGGAATATCTCAATTTCTTCCATGTTGTGTACAGCACAGT | 9360 |
| QY | 9361 | TCTGCTATGAAGTCTGCTTACAGAGAGAGATTAATAATCATATAGGATTAATAATCTTA | 9420 |
| Db | 9361 | TCTGCTATGAAGTCTGCTTACAGAGAGAGATTAATAATCATATAGGATTAATAATCTTA | 9420 |
| QY | 9421 | AGTTTGAAACAATGAGTTTTATAGTCAATTTGACATGAAGAAATTGAGACCTCTACTGG | 9480 |

| | | | |
|---|-------|--|-------|
| D | 9421 | AGTTTGAAGACATGAGGCTTTTACGTGCAATTTGACATGAAAGAAATGAGACCTCTACTGG | 9480 |
| Q | 9481 | ATAGCTATGGATTTATACGTCTCTTTTTCCTAGATTTACTATTGACCCACGTGAGTCAA | 9540 |
| D | 9481 | ATAGCTATGGATTTATACGTCTCTTTTTCCTAGATTTACTATTGACCCACGTGAGTCAA | 9540 |
| Q | 9541 | GTATGAACCTCAGCTCTCGGGCTACTGGCATTTGATTTACATACATCTGTAATTTTA | 9600 |
| D | 9541 | GTATGAACCTCAGCTCTCGGGCTACTGGCATTTGATTTACATACATCTGTAATTTTA | 9600 |
| Q | 9601 | GCAGGATTTTAGGGTTTATGAGTACTTTTGGAGTAATCATAGGCTTAAGTAATCTTAATC | 9660 |
| D | 9601 | GCAGGATTTTAGGGTTTATGAGTACTTTTGGAGTAATCATAGGCTTAAGTAATCTTAATC | 9660 |
| Q | 9661 | TCAGGGAAGAAAAAAGAACCAACCCGTGACAGACATCCAGCTCAGGTGGAAATCAAG | 9720 |
| D | 9661 | TCAGGGAAGAAAAAAGAACCAACCCGTGACAGACATCCAGCTCAGGTGGAAATCAAG | 9720 |
| Q | 9721 | ATCAGAGCTCAGTGGCGTCCAGAGAACAGGGACTTCTCTTTAGAGACCTTTATGTAC | 9780 |
| D | 9721 | ATCAGAGCTCAGTGGCGTCCAGAGAACAGGGACTTCTCTTTAGAGACCTTTATGTAC | 9780 |
| Q | 9781 | AGGGCTCAAGATTAATCTATGTATGTATGAGAGACTTTTCATTTCTGGCCACAGTTCACTG | 9840 |
| D | 9781 | AGGGCTCAAGATTAATCTATGTATGTATGAGAGACTTTTCATTTCTGGCCACAGTTCACTG | 9840 |
| Q | 9841 | AGGCATCTGGAAATTTTCTCCGCTCAGACAGTTCCAGTATCCAGTTTGTACAGTTTC | 9900 |
| D | 9841 | AGGCATCTGGAAATTTTCTCCGCTCAGACAGTTCCAGTATCCAGTTTGTACAGTTTC | 9900 |
| Q | 9901 | TGGCATTTTTGGCTCAGGGCCGTGATCCAGAGGACAGAAAGTTCCAGTATGTCTAGGGAG | 9960 |
| D | 9901 | TGGCATTTTTGGCTCAGGGCCGTGATCCAGAGGACAGAAAGTTCCAGTATGTCTAGGGAG | 9960 |
| Q | 9961 | TGGCTGACCGTCCACACTCATCTGCATCTAAACAAAGGGGAAACCAAGAAGTGGCTTTTG | 10020 |
| D | 9961 | TGGCTGACCGTCCACACTCATCTGCATCTAAACAAAGGGGAAACCAAGAAGTGGCTTTTG | 10020 |
| Q | 10021 | TTGAAATTTGCAGTGGGCCAGAGGGGCTGACCACTGATGATTTGACCAAGGACAA | 10080 |
| D | 10021 | TTGAAATTTGCAGTGGGCCAGAGGGGCTGACCACTGATGATTTGACCAAGGACAA | 10080 |
| Q | 10081 | TTAATCTCAGCAAGTGCATTTGCAAGCATTTAAATTTGAATCACTGATCAATGACA | 10140 |
| D | 10081 | TTAATCTCAGCAAGTGCATTTGCAAGCATTTAAATTTGAATCACTGATCAATGACA | 10140 |
| Q | 10141 | ATCACTATCAACAAGTGGTTGGCTTGGTAAGATGAGCTTAAGGGGCTCAAGAGTAGC | 10200 |
| D | 10141 | ATCACTATCAACAAGTGGTTGGCTTGGTAAGATGAGCTTAAGGGGCTCAAGAGTAGC | 10200 |
| Q | 10201 | TACTCTTAATGAGATTTGCATTTTGAAGCAGACACTGTGAAAAGCTGGCTCCTTAAGA | 10260 |
| D | 10201 | TACTCTCTAATGAGATTTGCATTTTGAAGCAGACACTGTGAAAAGCTGGCTCCTTAAGA | 10260 |
| Q | 10261 | GGCTGCTAAACATTAGGGTCAATTTTCCAGTGCACCTTTGTAAGTGTCTGCAGTTCCCA | 10320 |
| D | 10261 | GGCTGCTAAACATTAGGGTCAATTTTCCAGTGCACCTTTGTAAGTGTCTGCAGTTCCCA | 10320 |
| Q | 10321 | TGCAAGCTGCCCCAACATACACTTCCATTTGAATACATTTATATGACAGGCTACTGCT | 10380 |
| D | 10321 | TGCAAGCTGCCCCAACATACACTTCCATTTGAATACATTTATATGACAGGCTACTGCT | 10380 |
| Q | 10381 | TCTTTCGAGCACTGCTCTTCAANTGAACCAACCAAAATTTCAANAAGCTAGTGAAGA | 10440 |
| D | 10381 | TCTTTCGAGCACTGCTCTTCAANTGAACCAACCAAAATTTCAANAAGCTAGTGAAGA | 10440 |
| Q | 10441 | GTAAACAACCTTTGATATGATCAATTAAGTAATATATGCTTTCAGTAGTTCAGCTATTTAT | 10500 |
| D | 10441 | GTAAACAACCTTTGATATGATCAATTAAGTAATATATGCTTTCAGTAGTTCAGCTATTTAT | 10500 |
| Q | 10501 | GCCTCATGGAACAATCTTTGTACAAGCTGAACAATGCGGGCTTCAGATTTAGTGTAAAACCT | 10560 |

| | | | |
|----|-------|--|-------|
| Dh | 10501 | GCACACGTGAAGAACATCTTGTAACACCTGAACACTGGGGGCTCCAGATTAGTGGTAAACCT | 10506 |
| Qy | 10501 | ACTTTATACATCATAGATAATCATAGATAGGCTGGGTTGGAGAGGACCCCAAGATCATG | 10620 |
| Dh | 10561 | ACTTTATACATCATAGATAATCATAGATAGGCTGGGTTGGAGAGGACCCCAAGATCATG | 10620 |
| Qy | 10621 | AAGATCCAAACCCCGGACAGGAGGAGGACCAACCTCCAGATCTGGTACTACACAG | 10680 |
| Dh | 10621 | AAGATCCAAACCCCGGACAGGAGGAGGACCAACCTCCAGATCTGGTACTACACAG | 10680 |
| Qy | 10681 | GCAGCCCAAGGGGCTCATCCACACTGGGCATGAAACCTCCAGAGGATGAGCATCCACAAC | 10740 |
| Dh | 10681 | GCAGCCCAAGGGGCTCATCCACACTGGGCATGAAACCTCCAGAGGATGAGCATCCACAAC | 10740 |
| Qy | 10741 | CTCTGTGGGACGCTGTGGCCAGACCTCAGCACCTCTCTGTGAAGAACTTTTCCCTGAC | 10800 |
| Dh | 10741 | CTCTGTGGGACGCTGTGGCCAGACCTCAGCACCTCTCTGTGAAGAACTTTTCCCTGAC | 10800 |
| Qy | 10801 | ATCCAAATCAAGCCCTTCCCTCCTGTAGGTTAGATCCACTCCCCCTTGTGCTATCACTGTC | 10860 |
| Dh | 10801 | ATCCAAATCAAGCCCTTCCCTCCTGTAGGTTAGATCCACTCCCCCTTGTGCTATCACTGTC | 10860 |
| Qy | 10861 | TACTCTTGAAAAAATTGATTTCTCCCTCTTTTGGAAAGTTGCATGAGAGTCTCTTGCA | 10920 |
| Dh | 10861 | TACTCTTGAAAAAATTGATTTCTCCCTCTTTTGGAAAGTTGCATGAGAGTCTCTTGCA | 10920 |
| Qy | 10921 | GCCCTTCTCTCTCTCGACAGATGAACAAGCCAGCTCCCTCAGCCTGCTTTATAGAGA | 10980 |
| Dh | 10921 | GCCCTTCTCTCTCTCGACAGATGAACAAGCCAGCTCCCTCAGCCTGCTTTATAGAGA | 10980 |
| Qy | 10981 | GSTGCTCCAGCCCTCTGATCATCTTTTGTGGCCCTCCTCTGAGACCCGGTCCAGAGCTGCA | 11040 |
| Dh | 10981 | GSTGCTCCAGCCCTCTGATCATCTTTTGTGGCCCTCCTCTGAGACCCGGTCCAGAGCTGCA | 11040 |
| Qy | 11041 | CATCTTCTCTGACTGGGGGGCCCAAGGCGCTGATGCAGTACCTCAGATGGGAGCTCAAAA | 11100 |
| Dh | 11041 | CATCTTCTCTGACTGGGGGGCCCAAGGCGCTGATGCAGTACCTCAGATGGGAGCTCAAAA | 11100 |
| Qy | 11041 | CATCTTCTCTGACTGGGGGGCCCAAGGCGCTGATGCAGTACCTCAGATGGGAGCTCAAAA | 11100 |
| Dh | 11041 | CATCTTCTCTGACTGGGGGGCCCAAGGCGCTGATGCAGTACCTCAGATGGGAGCTCAAAA | 11100 |
| Qy | 11101 | GAGCAGAGTAAAGAGGGACATACCTTCTCACCCTGCTGGCCAGCCCTTCTGATGG | 11160 |
| Dh | 11101 | GAGCAGAGTAAAGAGGGACATACCTTCTCACCCTGCTGGCCAGCCCTTCTGATGG | 11160 |
| Qy | 11161 | AACCCCTGGATTAACACTGGCTTTCTAGCGTGAACCTTCTCCTATGACTGCACATTAATA | 11220 |
| Dh | 11161 | AACCCCTGGATTAACACTGGCTTTCTAGCGTGAACCTTCTCCTATGACTGCACATTAATA | 11220 |
| Qy | 11221 | ACAGGAACATATACAACAGTGTGATAGGCGACATGAGATTTTTCACACTTCTCAATTTG | 11280 |
| Dh | 11221 | ACAGGAACATATACAACAGTGTGATAGGCGACATGAGATTTTTCACACTTCTCAATTTG | 11280 |
| Qy | 11281 | GCTAGATCTTATAGTANGAGAACCTGTAAGTTGCTTCTGCTGCGTGTCTTCCCTCCAA | 11340 |
| Dh | 11281 | GCTAGATCTTATAGTANGAGAACCTGTAAGTTGCTTCTGCTGCGTGTCTTCCCTCCAA | 11340 |
| Qy | 11341 | ATATCTCTGCTGATATACCTCACCCCACTGCGCACTGATAGGCTCCATGGCCCCCTGACG | 11400 |
| Dh | 11341 | ATATCTCTGCTGATATACCTCACCCCACTGCGCACTGATAGGCTCCATGGCCCCCTGACG | 11400 |
| Qy | 11401 | CAGGGCCCTGATGAACCCGGGACCTGCTTGAAGTGGTTAATAGACAGTATGACCAAG | 11460 |
| Dh | 11401 | CAGGGCCCTGATGAACCCGGGACCTGCTTGAAGTGGTTAATAGACAGTATGACCAAG | 11460 |
| Qy | 11461 | TTGCACTATGTAATACACAACAAATGTGTTGGATCTCTTCAGCATGTGAGAACAGAGCA | 11520 |
| Dh | 11461 | TTGCACTATGTAATACACAACAAATGTGTTGGATCTCTTCAGCATGTGAGAACAGAGCA | 11520 |
| Qy | 11521 | AATTTGCATTTGCAGGAAATGGTTTAGTAATTTCTGCCAATTAAACTTGTATTATCACA | 11580 |
| Dh | 11521 | AATTTGCATTTGCAGGAAATGGTTTAGTAATTTCTGCCAATTAAACTTGTATTATCACA | 11580 |
| Qy | 11581 | TGGGCTTTTANAGCTGTATATAGTGTGACACTGATGATGAACAATGGCTATGCAAGTAA | 11640 |
| Dh | 11581 | TGGGCTTTTANAGCTGTATATAGTGTGACACTGATGATGAACAATGGCTATGCAAGTAA | 11640 |

| QY | 11641 | AATCAAGCCTGTAGTATTTGCAACAGACTATAAAATTCCTGTGGCTTGCCAAATGG | 11700 |
|-----------------------------|--|---|----------------------------|
| Db | 11641 | AATCAAGCCTGTAGTATTTGCAACAGACTATAAAATTCCTGTGGCTTGCCAAATGG | 11700 |
| QY | 11701 | TACTTCCCACTTGTATTAAGAAATTTGGCAAGTTTAAAGCAATTTGAAGTGTGGAA | 11760 |
| Db | 11701 | TACTTCCCACTTGTATTAAGAAATTTGGCAAGTTTAAAGCAATTTGAAGTGTGGAA | 11760 |
| QY | 11761 | ATTCTGTATTAAGCAAGGCGCTTTTGGACACTGTGTAAGACAGAGAATCAAAAGGGG | 11820 |
| Db | 11761 | ATTCTGTATTAAGCAAGGCGCTTTTGGACACTGTGTAAGACAGAGAATCAAAAGGGG | 11820 |
| QY | 11821 | TGGGAGCAAGTTAAAGAAAGGAGCGAGGCAAGAGACCTGCAGTCCCGCTGTGTACG | 11880 |
| Db | 11821 | TGGGAGCAAGTTAAAGAAAGGAGCGAGGCAAGAGACCTGCAGTCCCGCTGTGTACG | 11880 |
| QY | 11881 | ACACTGGCAACATGAGTCTTTGGCTAATCTGTTGCTTTCCTCCCGCTGGCTGCT | 11940 |
| Db | 11881 | ACACTGGCAACATGAGTCTTTGGCTAATCTGTTGCTTTCCTCCCGCTGGCTGCT | 11940 |
| QY | 11941 | TAGGG 11945 | |
| Db | 11941 | TAGGG 11945 | |
| RESULT 2 | AF405538 | 12728 bp | DNA linear SYN 06-SEP-2001 |
| LOCUS | AF405538 | | |
| DEFINITION | Synthetic construct lysozyme signal peptide/modified human interferon alpha 2b fusion protein precursor, gene, complete cds. | | |
| ACCESSION | AF405538 | | |
| VERSION | AF405538.1 | GI:15487987 | |
| KEYWORDS | | | |
| SOURCE | synthetic construct. | | |
| ORGANISM | synthetic construct | | |
| REFERENCE | 1 (bases 1 to 12728) | | |
| AUTHORS | Rapp, J.C. | | |
| TITLE | Chicken Lysozyme Promoter | | |
| JOURNAL | Unpublished | | |
| REFERENCE | 2 (bases 1 to 12728) | | |
| AUTHORS | Rapp, J.C. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (06-AUG-2001) Avigenics Inc, 425 River Road, Athens, Georgia 30602-2771, USA | | |
| FEATURES | Location/Qualifiers | | |
| Source | 1..12728 | | |
| | /organism="synthetic construct" | | |
| | /db_xref="taxon:32630" | | |
| | /note="cloning vector PAVICR-A115.93.1.2" | | |
| CDS | 11892..12443 | | |
| | /codon_start=1 | | |
| | /transl_table=1 | | |
| | /product="lysozyme signal peptide/modified human interferon alpha 2b fusion protein precursor" | | |
| | /protein_id="AA01039.1" | | |
| | /db_xref="GI:15487988" | | |
| | /translation="MRSLLIVLCPLIALGDIQPTHSLSRPTMLLAQWRRLSTFSCAKDHDFEPOEERGNQFQKATETPVILHEMIQOIFNLFTRKSSAANDETLDFVTELYQQINDIEACVIGVGVTETPLPKEDSILAVRKYFORITLYLKEKKYSPCAWEVRAEIMRFSSTMLQESLSRKE" | | |
| | 11892..11939 | | |
| | /note="derived from chicken lysozyme" | | |
| BASE COUNT | 3337 a 2755 c 2850 g 3786 t | | |
| ORIGIN | | | |
| Query Match | 100.0%; Score 11945; DB 12; Length 12728; | | |
| Best Local Similarity | 100.0%; Pred. No. 0; | | |
| Matches 11945; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; | | |
| QY | 1 | TCGCCCTCTTGTGATATTCACCTGTGTTGTAATTCCTCTTGTGCCGATGAAGATA | 60 |

Db 1 TCCGCCCTTCTTGATATTCACCTGTGTGTAATTCATCTCTTCGCCGATGAAGATA 60
Qy 61 TAACAGTCGTATTAACAGTCGTGTGAGAAATACCTGTATTTCTTCATCAGTCTTTT 120
Db 61 TAACAGTCGTATTAACAGTCGTGTGAGAAATACCTGTATTTCTTCATCAGTCTTTT 120
Qy 121 ATAAATATGTTGAATATTGGATTAAGCGTGTGTCTTTGTCTTGGAGACAAAGCCCA 180
Db 121 ATAAATATGTTGAATATTGGATTAAGCGTGTGTCTTTGTCTTGGAGACAAAGCCCA 180
Qy 181 CAGCAGGTGTGTGGGTGTGGCAGCTCAGTACAGAGATGTTTTTGGCTGT 240
Db 181 CAGCAGGTGTGTGGGTGTGGCAGCTCAGTACAGAGATGTTTTTGGCTGT 240
Qy 241 TTTTTTTTTTTTTTTTTTAAAGGTGTCTTTTCTTGTAAATTTTCTACTGA 300
Db 241 TTTTTTTTTTTTTTTTTTAAAGGTGTCTTTTCTTGTAAATTTTCTACTGA 300
Qy 301 CTGTATGTTTGTACAGGTGTACAGAAACATTTCTCAAAAGAGACCTTTTGAACCTGTAC 360
Db 301 CTGTATGTTTGTACAGGTGTACAGAAACATTTCTCAAAAGAGACCTTTTGAACCTGTAC 360
Qy 361 AGCCCTTTCTTTCATCTCCCTTTTGTCTTGTGCAATGCCCTTGGTGTGATGTGAT 420
Db 361 AGCCCTTTCTTTCATCTCCCTTTTGTCTTGTGCAATGCCCTTGGTGTGATGTGAT 420
Qy 421 TATGGAACACGTTGATGCGAACCTTGAGCTTTTATTTATAGTGTGCTGAAGCCTTGA 480
Db 421 TATGGAACACGTTGATGCGAACCTTGAGCTTTTATTTATAGTGTGCTGAAGCCTTGA 480
Qy 481 TAGCTGTGTACACAGATACCTTATTAAGTTAGGCCAGCTGTATGCTTATTTTTC 540
Db 481 TAGCTGTGTACACAGATACCTTATTAAGTTAGGCCAGCTGTATGCTTATTTTTC 540
Qy 541 CCTTTGAAGTAGTAGCTCTCTGTGTCTTTTCTTTGAACTGTGAGGCTTATGATTT 600
Db 541 CCTTTGAAGTAGTAGCTCTCTGTGTCTTTTCTTTGAACTGTGAGGCTTATGATTT 600
Qy 601 TTCTAATAGGATTTTTTACCTGATATCTAGTGCATACCCAAATGCTGTGAATGTTTT 660
Db 601 TTCTAATAGGATTTTTTACCTGATATCTAGTGCATACCCAAATGCTGTGAATGTTTT 660
Qy 661 CCTAGTAAACATGTTGATTAACCTGATTTACATGTTATATCTGTACATCTGTGT 720
Db 661 CCTAGTAAACATGTTGATTAACCTGATTTACATGTTATATCTGTACATCTGTGT 720
Qy 721 CTAGTAAATAATATAGCATTATAGAAATAGTAATCTGATTTCTTTTATATC 780
Db 721 CTAGTAAATAATATAGCATTATAGAAATAGTAATCTGATTTCTTTTATATC 780
Qy 781 TCTATGCTGTGTGTACAGTCAACAGACTTACCTCCTATTTTTTATTAAGATTTT 840
Db 781 TCTATGCTGTGTGTACAGTCAACAGACTTACCTCCTATTTTTTATTAAGATTTT 840
Qy 841 ATATGACATGTGTGTGTCTTGTGTAGGATACAGCTTAATTTCTTAAGACG 900
Db 841 ATATGACATGTGTGTGTCTTGTGTAGGATACAGCTTAATTTCTTAAGACG 900
Qy 901 ATGCTCAGTAAAGCGGTGTGTACATGGGTCAATGTAAACGGGCAAGTGTGGCTGT 960
Db 901 ATGCTCAGTAAAGCGGTGTGTACATGGGTCAATGTAAACGGGCAAGTGTGGCTGT 960
Qy 961 GCCTTCCCGAGATCCAGAGACCTAAACGCTTGTGACAGTAAATGCTTCAGAT 1020
Db 961 GCCTTCCCGAGATCCAGAGACCTAAACGCTTGTGACAGTAAATGCTTCAGAT 1020
Qy 1021 CCCAGGGAAGTGCAGATCCAGTGCATTTCTTAAGGAAGATGAATCTTTCTAAATA 1080
Db 1021 CCCAGGGAAGTGCAGATCCAGTGCATTTCTTAAGGAAGATGAATCTTTCTAAATA 1080
Qy 1081 TTTTGGCATAGGAAGCAAGCTGATGATTTGTTGGGACTTAAATTTTGGTAACGG 1140
Db 1081 TTTTGGCATAGGAAGCAAGCTGATGATTTGTTGGGACTTAAATTTTGGTAACGG 1140

Qy 1141 AGTCATAGGTTTTTAAACAGATTGCAGATGCTAACAGTACACAGCTTTATGCAAG 1200
Db 1141 AGTCATAGGTTTTTAAACAGATTGCAGATGCTAACAGTACACAGCTTTATGCAAG 1200
Qy 1201 TGATGCTTGATGCTGTGTCAGCTGTTTACGGCAGCTCCTTTCAGTAGAGATTCAGAT 1260
Db 1201 TGATGCTTGATGCTGTGTCAGCTGTTTACGGCAGCTCCTTTCAGTAGAGATTCAGAT 1260
Qy 1261 AGGGTGGGTCCTTGTGTGTGTCTCCACAGCTGTCCACAGCCACCTCCGGACA 1320
Db 1261 AGGGTGGGTCCTTGTGTGTGTCTCCACAGCTGTCCACAGCCACCTCCGGACA 1320
Qy 1321 CATCTCACCTGTGGGTACTTTTCAACCATTTTGCAGTAGATAGTACTATGAA 1380
Db 1321 CATCTCACCTGTGGGTACTTTTCAACCATTTTGCAGTAGATAGTACTATGAA 1380
Qy 1381 ACAGGAAGTTCCTAGTGGATATTCTCATGAGATGTCCTTTTCCATGTGGGCAAA 1440
Db 1381 ACAGGAAGTTCCTAGTGGATATTCTCATGAGATGTCCTTTTCCATGTGGGCAAA 1440
Qy 1441 GTATGATTAAGCATCTATTTGTAAATTTATGCACTGTGTGTCGTAATCTTCTAT 1500
Db 1441 GTATGATTAAGCATCTATTTGTAAATTTATGCACTGTGTGTCGTAATCTTCTAT 1500
Qy 1501 AGCACCACTTATTGACAGAGGTATAGGCTGTGTGTGCTGTGTGCTTCAATCTT 1560
Db 1501 AGCACCACTTATTGACAGAGGTATAGGCTGTGTGTGCTGTGTGCTTCAATCTT 1560
Qy 1561 TTAAGCTTTTGGAAATACACGTACTGATGTGAAGTCTTTGAAGATGTAAACAGTA 1620
Db 1561 TTAAGCTTTTGGAAATACACGTACTGATGTGAAGTCTTTGAAGATGTAAACAGTA 1620
Qy 1621 CTATCCTTTGATCCCAATGAATGAGCATTCAGTTGTAAAGATCCCTCATTCAT 1680
Db 1621 CTATCCTTTGATCCCAATGAATGAGCATTCAGTTGTAAAGATCCCTCATTCAT 1680
Qy 1681 ACCATGATATGATTTTACACCCCACTGCTGACACTTGTGAATATTCAAGTAATAG 1740
Db 1681 ACCATGATATGATTTTACACCCCACTGCTGACACTTGTGAATATTCAAGTAATAG 1740
Qy 1741 ACTTTGGCTCACCCCTGTGTGTGTGTGTGTGTGTGTGAATGAAGAAATTTTAACTGTCA 1800
Db 1741 ACTTTGGCTCACCCCTGTGTGTGTGTGTGTGTGTGTGAATGAAGAAATTTTAACTGTCA 1800
Qy 1801 TATGATTTATCATTTATGAAGAGACATTCGCTGATCTTCAANTGTAAAGAAATGAGGA 1860
Db 1801 TATGATTTATCATTTATGAAGAGACATTCGCTGATCTTCAANTGTAAAGAAATGAGGA 1860
Qy 1861 GTGGGTGTCTTTTATTAATACAAAGTGAATCAAAATAGTSCAGGTGTCTTAAAAAAA 1920
Db 1861 GTGGGTGTCTTTTATTAATACAAAGTGAATCAAAATAGTSCAGGTGTCTTAAAAAAA 1920
Qy 1921 AAAAAAAGTAAATTAAGAGACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Db 1921 AAAAAAAGTAAATTAAGAGACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
Qy 1981 TAAACAGTTTACATTTTATGAAGATTAACAGCTGTGCTGATTTCTTAAACATTAAGGCTGT 2040
Db 1981 TAAACAGTTTACATTTTATGAAGATTAACAGCTGTGCTGATTTCTTAAACATTAAGGCTGT 2040
Qy 2041 ATTTGCTTCCCTGTACCATGTCATTTTCCATTTCCCAATTTTSCAACAAGATGTGTGGTAA 2100
Db 2041 ATTTGCTTCCCTGTACCATGTCATTTTCCATTTCCCAATTTTSCAACAAGATGTGTGGTAA 2100
Qy 2101 ACTATTCAGAAATGCTTTTAAATACAGATGGAGCTGTCTGATTTGGAATGAGAG 2160
Db 2101 ACTATTCAGAAATGCTTTTAAATACAGATGGAGCTGTCTGATTTGGAATGAGAG 2160
Qy 2161 TTGCACCTCAAAATGTACAGAAATGATGTCTCAGAAATGCCCAACTCCAAAGGATTTT 2220
Db 2161 TTGCACCTCAAAATGTACAGAAATGATGTCTCAGAAATGCCCAACTCCAAAGGATTTT 2220

| | | | |
|----|------|--|------|
| Qy | 2221 | ATATGCTATATATGTAAGCAAGTTTCCTGATTTCCAGACAGGCCAAAGATCTCTGAATGT | 2280 |
| Db | 2221 | ATATGCTATATATGTAAGCAAGTTTCCTGATTTCCAGACAGGCCAAAGATCTCTGAAGTT | 2280 |
| Qy | 2281 | GTTGTGCGGAGACCTGTATTTCTCAACAAGGTAAAGATGATTCCTACCACTGCGATTT | 2340 |
| Db | 2281 | GTTGTGCGGAGACCTGTATTTCTCAACAAGGTAAAGATGATTCCTACCACTGCGATTT | 2340 |
| Qy | 2341 | TTAATACATTTTCGACGAGAAGTACTAGTTAAATCTCTACCTTTAGGGATGCTTCATCAT | 2400 |
| Db | 2341 | TTAATACATTTTCGACGAGAAGTACTAGTTAAATCTCTACCTTTAGGGATGCTTCATCAT | 2400 |
| Qy | 2401 | TTTTAGATGTTATATACCTTGAATATCTGACATATCTTTAGCTTTCATGCGGTCCCTTTTTC | 2460 |
| Db | 2401 | TTTTAGATGTTATATACCTTGAATATCTGACATATCTTTAGCTTTCATGCGGTCCCTTTTTC | 2460 |
| Qy | 2461 | AGCCTTTAGAGACCTGTTAAGCAATTTGCTGTCCAACTTTGTGTGGCTTTAAACTGCA | 2520 |
| Db | 2461 | AGCCTTTAGAGACCTGTTAAGCAATTTGCTGTCCAACTTTGTGTGGCTTTAAACTGCA | 2520 |
| Qy | 2521 | ATATAGATTTACCTTGTATTTGAAGAAATAAAGACATTTTATATTTAAAAAATACCTTTG | 2580 |
| Db | 2521 | ATATAGATTTACCTTGTATTTGAAGAAATAAAGACATTTTATATTTAAAAAATACCTTTG | 2580 |
| Qy | 2581 | TCTGTCTTCATTTTGAAGTCTGTGATATCTTCGATATCTTCGACGTGCTGACG | 2640 |
| Db | 2581 | TCTGTCTTCATTTTGAAGTCTGTGATATCTTCGATATCTTCGACGTGCTGACG | 2640 |
| Qy | 2641 | ATATATCAGACATCATAAACTTAAACGTAGAGCTAGTGAGTTACAGCTGGGGTTTTGATGCT | 2700 |
| Db | 2641 | ATATATCAGACATCATAAACTTAAACGTAGAGCTAGTGAGTTACAGCTGGGGTTTTGATGCT | 2700 |
| Qy | 2701 | GTTATATATTTCTGAAACCTAGAAATGATGTTGTCTTCATCTGCTCATCAACACCTTCATGC | 2760 |
| Db | 2701 | GTTATATATTTCTGAAACCTAGAAATGATGTTGTCTTCATCTGCTCATCAACACCTTCATGC | 2760 |
| Qy | 2761 | AGAGTGAAGGCTATGAGAGAAATGCAATTTATTTGATTACTTTTAAAGTCAACTTTT | 2820 |
| Db | 2761 | AGAGTGAAGGCTATGAGAGAAATGCAATTTATTTGATTACTTTTAAAGTCAACTTTT | 2820 |
| Qy | 2821 | TATCAGATTTTTTTTTCATTTTGAATATATTTGTTTCTTACAGCTGCATAGCTTCGATTC | 2880 |
| Db | 2821 | TATCAGATTTTTTTTTCATTTTGAATATATTTGTTTCTTACAGCTGCATAGCTTCGATTC | 2880 |
| Qy | 2881 | TGAATGTACATCTGATTTGSCATGAGAGACACAGCACTTTCATCTTAACTTAACTTCAT | 2940 |
| Db | 2881 | TGAATGTACATCTGATTTGSCATGAGAGACACAGCACTTTCATCTTAACTTAACTTCAT | 2940 |
| Qy | 2941 | TTTGTGAATGAAGGATTTAAGCAGAGGGCACAGGCTCATGAATATAGACACGTGGCTCAG | 3000 |
| Db | 2941 | TTTGTGAATGAAGGATTTAAGCAGAGGGCACAGGCTCATGAATATAGACACGTGGCTCAG | 3000 |
| Qy | 3001 | GAGAAATGGAACCGGATTTCTTGGCTAGTGCTGCTAAATCTGTGTAGAGAAAGTAAACA | 3060 |
| Db | 3001 | GAGAAATGGAACCGGATTTCTTGGCTAGTGCTGCTAAATCTGTGTAGAGAAAGTAAACA | 3060 |
| Qy | 3061 | CCCGATTTCTTGAAGAGGGCTCCACGCTTTTATGCTTCCAAATTTGAAGGTGGCAGGCACTT | 3120 |
| Db | 3061 | CCCGATTTCTTGAAGAGGGCTCCACGCTTTTATGCTTCCAAATTTGAAGGTGGCAGGCACTT | 3120 |
| Qy | 3121 | GGCCACTGTGTTATTTACTGATATATGTCTCAGTTTTCGACGCTAACCTGTGCTTCCACTA | 3180 |
| Db | 3121 | GGCCACTGTGTTATTTACTGATATATGTCTCAGTTTTCGACGCTAACCTGTGCTTCCACTA | 3180 |
| Qy | 3181 | TTTGAGCATGACATTAATGCTTCGCTTGAAGGCGAGGTGAAGGTTGGATGGGTGGGAAGA | 3240 |
| Db | 3181 | TTTGAGCATGACATTAATGCTTCGCTTGAAGGCGAGGTGAAGGTTGGATGGGTGGGAAGA | 3240 |
| Qy | 3241 | GTTGCTGGGCTGTGGCTGGGGGAGCTGTGGGAGCTCCAACTTAGCTTGGGGTGGGACAGA | 3300 |
| Db | 3241 | GTTGCTGGGCTGTGGCTGGGGGAGCTGTGGGAGCTCCAACTTAGCTTGGGGTGGGACAGA | 3300 |
| Qy | 3301 | CAGGGAAGAAGTGTGCTAACTATTTTAAAGTACTGTGTTCGCAACGTCATCTGCAAAAT | 3360 |

| | | | |
|----|------|---|------|
| Dd | 3301 | CAGGGAAAGTGTGGGTACTATTTTAAAGTCTGTGGCAAGCTCTCATCTGCAAT | 3360 |
| Qy | 3361 | ACGTAGGGTGTCTACTCTCGAAGATTAAACAGTGTGGTTCCAGTAATATATGATGAATTC | 3420 |
| Dd | 3361 | ACGTAGGGGTGTCTACTCTCGAAGATTAAACAGTGTGGTTCCAGTAATATATGATGAATTC | 3420 |
| Qy | 3421 | ACAGTGAAGCATTCGAAGGGTAGATCATCTAAACGACCAGATCATCAAGCTATGATTTGG | 3480 |
| Dd | 3421 | ACAGTGAAGCATTCGAAGGGTAGATCATCTAAACGACCAGATCATCAAGCTATGATTTGG | 3480 |
| Qy | 3481 | AAGCGGTATCAAAAGGAGGAGTAAGCAAGCTTCATATGTTTCCCTCCACGTAA | 3540 |
| Dd | 3481 | AAGCGGTATCAAAAGGAGGAGTAAGCAAGCTTCATATGTTTCCCTCCACGTAA | 3540 |
| Qy | 3541 | GCAGTCTGGGAAGTAGACACCCCTTGACAGACAGCAAGAAATATTCAGAGCATGTGC | 3600 |
| Dd | 3541 | GCAGTCTGGGAAGTAGACACCCCTTGACAGACAGCAAGAAATATTCAGAGCATGTGC | 3600 |
| Qy | 3601 | TAGGAACAATTCTTGTCTGTAATTCATCTGCAAGACTTTGATGCTGCTGTGTGGCC | 3660 |
| Dd | 3601 | TAGGAACAATTCTTGTCTGTAATTCATCTGCAAGACTTTGATGCTGCTGTGTGGCC | 3660 |
| Qy | 3661 | TTCTCGAGACCTCGAAGGCCGAGAGCCGTGTGTAGCTGGAGGGAAAGATTTCGCTAA | 3720 |
| Dd | 3661 | TTCTCGAGACCTCGAAGGCCGAGAGCCGTGTGTAGCTGGAGGGAAAGATTTCGCTAA | 3720 |
| Qy | 3721 | GTCCAAAGCTTCAGCAGGTATTTGTCTTTCCTTCCCCAGCACTGTGCAGCAGATGG | 3780 |
| Dd | 3721 | GTCCAAAGCTTCAGCAGGTATTTGTCTTTCCTTCCCCAGCACTGTGCAGCAGATGG | 3780 |
| Qy | 3781 | AACTGATGTGCAAGCCTCTCTGTCCACTACCTGTCTCGAAGCAGACTGCTCGAATA | 3840 |
| Dd | 3781 | AACTGATGTGCAAGCCTCTCTGTCCACTACCTGTCTCGAAGCAGACTGCTCGAATA | 3840 |
| Qy | 3841 | AAGAGCTTACTCTATGCGATAGTGTGAAAGTAATGGTTTAAAAAAGAAACACAA | 3900 |
| Dd | 3841 | AAGAGCTTACTCTATGCGATAGTGTGAAAGTAATGGTTTAAAAAAGAAACACAA | 3900 |
| Qy | 3901 | AGGCAAAACCGGCTCCGCCATGAGAAGAAAGCAGTGTAAACATGATGAAAGGTGCAG | 3960 |
| Dd | 3901 | AGGCAAAACCGGCTCCGCCATGAGAAGAAAGCAGTGTAAACATGATGAAAGGTGCAG | 3960 |
| Qy | 3961 | AAGCCCCAGGAGGTGTGACAGGCCCTTCCTCCACCTAGAGCGGGAACAACGTTCCCT | 4020 |
| Dd | 3961 | AAGCCCCAGGAGGTGTGACAGGCCCTTCCTCCACCTAGAGCGGGAACAACGTTCCCT | 4020 |
| Qy | 4021 | GGCTAGGGCTCTGCCCCGGAAGTGGCTGTTCTTTGTGTGGTTTGTGGCTTTGGTT | 4080 |
| Dd | 4021 | GGCTAGGGCTCTGCCCCGGAAGTGGCTGTTCTTTGTGTGGTTTGTGGCTTTGGTT | 4080 |
| Qy | 4081 | TTTGAGTTTATGACAAAGGAAGCCTGAAAGAGGTGTTGGGCACATATTTTGGTTGTA | 4140 |
| Dd | 4081 | TTTGAGTTTATGACAAAGGAAGCCTGAAAGAGGTGTTGGGCACATATTTTGGTTGTA | 4140 |
| Qy | 4141 | AACCTGTACTTAAATATATATTTTGTGAGGAGTGTAGCGAATTTGGCCAAATTTAAATA | 4200 |
| Dd | 4141 | AACCTGTACTTAAATATATATTTTGTGAGGAGTGTAGCGAATTTGGCCAAATTTAAATA | 4200 |
| Qy | 4201 | AAGTTGCAAGACATTGAAGCTGAGTGTGAGAGGTAACACGTTTAAATGAGATCTTCT | 4260 |
| Dd | 4201 | AAGTTGCAAGACATTGAAGCTGAGTGTGAGAGGTAACACGTTTAAATGAGATCTTCT | 4260 |
| Qy | 4261 | GAAACTACTGCTTAAACACTGTTTGAAGTGTAGACCTTGGATAGGTAGTGTCTTT | 4320 |
| Dd | 4261 | GAAACTACTGCTTAAACACTGTTTGAAGTGTAGACCTTGGATAGGTAGTGTCTTT | 4320 |
| Qy | 4321 | GTTACATGTCTGATACACTTGTCTTCCATTCCATCCATCCATTCGATTCGACATCCA | 4380 |
| Dd | 4321 | GTTACATGTCTGATACACTTGTCTTCCATTCCATCCATCCATTCGATTCGACATCCA | 4380 |
| Qy | 4381 | GCGATTTGTCACTTATCCCAATATCTGTCAATATCTGACATACCTGTCTCTGTCACATTGG | 4440 |
| Dd | 4381 | GCGATTTGTCACTTATCCCAATATCTGTCAATATCTGACATACCTGTCTCTGTCACATTGG | 4440 |

Db 4381 CGCATTTGCTACTTATCCCATATCTGTCAATCTGACATACCTGTCTCTTCTGCTACTTGG 4440
Qy 4441 TCAGAAGAAGAGATGTATATAATCCCAAGCCGCCCAAGTTTGAGAAGATGGCAGTTGCT 4500
Db 4441 TCAGAAGAAGAGATGTATATAATCCCAAGCCGCCCAAGTTTGAGAAGATGGCAGTTGCT 4500
Qy 4501 TCTTCCCTTTTCTGCTAGTAAAGATTTTCTCTGCTTTGACACTGACGAATAG 4560
Db 4501 TCTTCCCTTTTCTGCTAGTAAAGATTTTCTCTGCTTTGACACTGACGAATAG 4560
Qy 4561 TCTTCCCTTTTCTGCTAGTAAAGATTTTCTCTGCTTTGACACTGACGAATAG 4620
Db 4561 TCTTCCCTTTTCTGCTAGTAAAGATTTTCTCTGCTTTGACACTGACGAATAG 4620
Qy 4621 TTTGTGCTTCTCTACTCTAGTAAAGATTTTCTCTGCTTTGACACTGACGAATAG 4680
Db 4621 TTTGTGCTTCTCTACTCTAGTAAAGATTTTCTCTGCTTTGACACTGACGAATAG 4680
Qy 4681 GTTGGCCGAGTCTCTGATGAACACCTGATATATGGCCAAAGGTGGGTGGTTC 4740
Db 4681 GTTGGCCGAGTCTCTGATGAACACCTGATATATGGCCAAAGGTGGGTGGTTC 4740
Qy 4741 TCTGAGAACGGGACGCTTTTGGCTCTGAAAGCAAGAGACTCTGGGAGTTGCAATTAT 4800
Db 4741 TCTGAGAACGGGACGCTTTTGGCTCTGAAAGCAAGAGACTCTGGGAGTTGCAATTAT 4800
Qy 4801 TTGCAACTGATGTGGAACTGCTGCTTAAGCAGATTCCTAGTTCCTGCTACTTCT 4860
Db 4801 TTGCAACTGATGTGGAACTGCTGCTTAAGCAGATTCCTAGTTCCTGCTACTTCT 4860
Qy 4861 TTCTCTTCTGAGTCACTTATTTCTGACAGACAAACAGCCACCCGACAGGCTTA 4920
Db 4861 TTCTCTTCTGAGTCACTTATTTCTGACAGACAAACAGCCACCCGACAGGCTTA 4920
Qy 4921 GAAAGTATGTGGCTCTGCTGGGTGTGTTACAGCTCTGCCCTGTAAGAGGGGATTA 4980
Db 4921 GAAAGTATGTGGCTCTGCTGGGTGTGTTACAGCTCTGCCCTGTAAGAGGGGATTA 4980
Qy 4981 CGGGACCACTTCATCCCAAGAGATCTCATTCATGATGATCAAGCTGTAAGAACTTGG 5040
Db 4981 CGGGACCACTTCATCCCAAGAGATCTCATTCATGATGATCAAGCTGTAAGAACTTGG 5040
Qy 5041 CTCCAACCTCAAAACATTAATTTGAGTGAATGAATGAATGAATGAATGAATGAAT 5100
Db 5041 CTCCAACCTCAAAACATTAATTTGAGTGAATGAATGAATGAATGAATGAATGAAT 5100
Qy 5101 TAAAGTATTTAGTCTGAGCTCTGACGATGTAGTGTGAGTGTGAGTGTGAGTGT 5160
Db 5101 TAAAGTATTTAGTCTGAGCTCTGACGATGTAGTGTGAGTGTGAGTGTGAGTGT 5160
Qy 5161 CACTGATGAGAGTATGAATAAATGAGAGACGATTCAGAACCAACGAGGTGTGCG 5220
Db 5161 CACTGATGAGAGTATGAATAAATGAGAGACGATTCAGAACCAACGAGGTGTGCG 5220
Qy 5221 AAGAAACGATGGAATTAATCATGATTTGTGTGTGACATTTTAAATACATATA 5280
Db 5221 AAGAAACGATGGAATTAATCATGATTTGTGTGTGACATTTTAAATACATATA 5280
Qy 5281 CTACTTCAATGAGTGGGAGAGTCAAGTGTATTAGACGCCATTAACCAAGTGA 5340
Db 5281 CTACTTCAATGAGTGGGAGAGTCAAGTGTATTAGACGCCATTAACCAAGTGA 5340
Qy 5341 CGACTACCATTTTCTCTACAGAAAAACGATTCGAGCTCTGGGTAGTATAGTCTC 5400
Db 5341 CGAGTACCATTTTCTCTACAGAAAAACGATTCGAGCTCTGGGTAGTATAGTCTC 5400
Qy 5401 CATAGCGGCTAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5460
Db 5401 CATAGCGGCTAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5460
Qy 5461 CTTGGGGTTTCTCTACAGAGATTAAGGACAACTACTTCAAAAAATTTCTTTTCC 5520
Db 5461 CTTGGGGTTTCTCTACAGAGATTAAGGACAACTACTTCAAAAAATTTCTTTTCC 5520

Qy 5521 TCTCATGTGGATTCCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580
Db 5521 TCTCATGTGGATTCCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580
Qy 5581 CAGCGGTTTGGAAAAAGAAATTTGGAATTAACATGTCTAGCTTACCTCT 5640
Db 5581 CAGCGGTTTGGAAAAAGAAATTTGGAATTAACATGTCTAGCTTACCTCT 5640
Qy 5641 CCAGCATTTTGGTTTATTAATGATTAACCTGCTAGATTTGGAATGAGAGGGT 5700
Db 5641 CCAGCATTTTGGTTTATTAATGATTAACCTGCTAGATTTGGAATGAGAGGGT 5700
Qy 5701 TGGGTATTAACGAGAACAAAGAGCTTATTAATTAACCTGCTTATTTAGAGA 5760
Db 5701 TGGGTATTAACGAGAACAAAGAGCTTATTAATTAACCTGCTTATTTAGAGA 5760
Qy 5761 ACTGCGAGCTGTCAAAAAAAGGCTTACCCCAATTAAGTGAATAGCCCTTA 5820
Db 5761 ACTGCGAGCTGTCAAAAAAAGGCTTACCCCAATTAAGTGAATAGCCCTTA 5820
Qy 5821 GCCACAGGCGCAGCAGAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5880
Db 5821 GCCACAGGCGCAGCAGAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5880
Qy 5881 ACTGAGAGCAACTGCTTTGGAATGACAGCACTTGTGCAATTCCTTTGTTGAGA 5940
Db 5881 ACTGAGAGCAACTGCTTTGGAATGACAGCACTTGTGCAATTCCTTTGTTGAGA 5940
Qy 5941 TGGGTAGAGGCTGCTTGGGAGAGTTCCTGATTAAGGCACTTCTTTTCTCTG 6000
Db 5941 TGGGTAGAGGCTGCTTGGGAGAGTTCCTGATTAAGGCACTTCTTTTCTCTG 6000
Qy 6001 TCTCATTCCTCTAAGATCTCTCATGCTGTAATCCAGTCAAGTGAAGCTTCAACA 6060
Db 6001 TCTCATTCCTCTAAGATCTCTCATGCTGTAATCCAGTCAAGTGAAGCTTCAACA 6060
Qy 6061 ATGAATCATGACGTAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6120
Db 6061 ATGAATCATGACGTAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6120
Qy 6121 ATGAAGCTTATTAATTTTGTGCTTCCATATCATGCTTCTATGAATTCATATC 6180
Db 6121 ATGAAGCTTATTAATTTTGTGCTTCCATATCATGCTTCTATGAATTCATATC 6180
Qy 6181 ACCACGCAATTAAGTGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6240
Db 6181 ACCACGCAATTAAGTGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6240
Qy 6241 TTTCTCTTGGCAAGGCAAGGAAATGCTGAGAGCTTGAATTAACGCTGGGGTAGA 6300
Db 6241 TTTCTCTTGGCAAGGCAAGGAAATGCTGAGAGCTTGAATTAACGCTGGGGTAGA 6300
Qy 6301 AGTTCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6360
Db 6301 AGTTCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6360
Qy 6361 AGGACCAATTAAGGCTTATCTGCGGTTTGTCTGCTGCTGCTGCTGCTGCT 6420
Db 6361 AGGACCAATTAAGGCTTATCTGCGGTTTGTCTGCTGCTGCTGCTGCTGCT 6420
Qy 6421 CACTATTTCACTGCTCCAGGCTTACAAACCAAGATTAACGCTGGAATTTTCTAG 6480
Db 6421 CACTATTTCACTGCTCCAGGCTTACAAACCAAGATTAACGCTGGAATTTTCTAG 6480
Qy 6481 ACATTAATTAATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6540
Db 6481 ACATTAATTAATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6540
Qy 6541 CTGTTTAAACCCCTTAAGGATTCAGAACCACTAGATCATAGAAATGTTGATTGGA 6600
Db 6541 CTGTTTAAACCCCTTAAGGATTCAGAACCACTAGATCATAGAAATGTTGATTGGA 6600

Db 8761 TTTTCTCTGCTGTCAGAACATTTTGATPACAGAGAAAAGAAAGTGTCTTCTT 8820
Qy 8821 GGCATGGAGAGAGTGTGTCACTTGCATAATAAGAGATGCAGTCCCAATGTTCATATC 8880
Db 8821 GGCATGGAGAGAGTGTGTCACTTGCATAATAAGAGATGCAGTCCCAATGTTCATATC 8880
Qy 8881 TCAGGGCTGAGAGAGATCAGAACTGTGTATACAAATTTTCAGAGCTTCTGTATATCAGC 8940
Db 8881 TCAGGGCTGAGAGAGATCAGAACTGTGTATACAAATTTTCAGAGCTTCTGTATATCAGC 8940
Qy 8941 TTTTGAAGCTGTCTGCGCCAGGAGTACTAGTGAAGACCTCGGAAACAGAGACAA 9000
Db 8941 TTTTGAAGCTGTCTGCGCCAGGAGTACTAGTGAAGACCTCGGAAACAGAGACAA 9000
Qy 9001 TGTCTTCAAGGTGTCAGCAGAGAGAAACCTTGTCCCATATGAAATGAATACACATGC 9060
Db 9001 TGTCTTCAAGGTGTCAGCAGAGAGAAACCTTGTCCCATATGAAATGAATACACATGC 9060
Qy 9061 GCCTGAAGAAATCCAGCTCTGTGTTGAGCAGGTGGTGCACCTCCACACTGAACACAA 9120
Db 9061 GCCTGAAGAAATCCAGCTCTGTGTTGAGCAGGTGGTGCACCTCCACACTGAACACAA 9120
Qy 9121 GTTCATTTTTTATAGACTTCCAGAGAGATCTTCTTAACTTCTTAAATATGTACA 9180
Db 9121 GTTCATTTTTTATAGACTTCCAGAGAGATCTTCTTAACTTCTTAAATATGTACA 9180
Qy 9181 TCTCCAGTTGGAGATGACTATGACTACTGACAGAGAAATGAGAACTAGCTGGAAAT 9240
Db 9181 TCTCCAGTTGGAGATGACTATGACTACTGACAGAGAAATGAGAACTAGCTGGAAAT 9240
Qy 9241 TTTCTTTTGACCACCATGGAGTCAACCATTTCTTACTGCTGTAATTTGAATATATTT 9300
Db 9241 TTTCTTTTGACCACCATGGAGTCAACCATTTCTTACTGCTGTAATTTGAATATATTT 9300
Qy 9301 GAATTTGCAAGCAGAGATGAGGAAATCTTCAATTTTGCATGTTGGTGACAGACAGT 9360
Db 9301 GAATTTGCAAGCAGAGATGAGGAAATCTTCAATTTTGCATGTTGGTGACAGACAGT 9360
Qy 9361 TCTGCTTTGAAGTCTGCTTACAGGAAAGAGATTAATAATATGAGGATTAATATCTA 9420
Db 9361 TCTGCTTTGAAGTCTGCTTACAGGAAAGAGATTAATAATATGAGGATTAATATCTA 9420
Qy 9421 AGTTTGAAGACAAATGAGTTTATGCTGATTTGACATGAGAAATTTGAGACCTTACTG 9480
Db 9421 AGTTTGAAGACAAATGAGTTTATGCTGATTTGACATGAGAAATTTGAGACCTTACTG 9480
Qy 9481 ATAGCTATGATTTTACGTCTCTTTTCTTACTTACTTATGACCCAGCTGAGGTCAA 9540
Db 9481 ATAGCTATGATTTTACGTCTCTTTTCTTACTTACTTATGACCCAGCTGAGGTCAA 9540
Qy 9541 GTATGAACCTAGGTCTCTCGGCTCTAGCATGATGATTGATTAACAATCTTAATTTTA 9600
Db 9541 GTATGAACCTAGGTCTCTCGGCTCTAGCATGATGATTGATTAACAATCTTAATTTTA 9600
Qy 9601 GCAGTATTTAGGGTTTATGATTAATTTTGTGCAATCATGAGGTTGATATGTAATC 9660
Db 9601 GCAGTATTTAGGGTTTATGATTAATTTTGTGCAATCATGAGGTTGATATGTAATC 9660
Qy 9661 TCAGGGAAGAAAAAAGCCAACTGACAGACATCCAGCTCAGGTGGAATCAAGG 9720
Db 9661 TCAGGGAAGAAAAAAGCCAACTGACAGACATCCAGCTCAGGTGGAATCAAGG 9720
Qy 9721 ATTCAGCTGAGGTGCTGCGCCAGAGAAACAGAGACTTCTCTTGAAGACTTTATGTAAC 9780
Db 9721 ATTCAGCTGAGGTGCTGCGCCAGAGAAACAGAGACTTCTCTTGAAGACTTTATGTAAC 9780
Qy 9781 AGGGCTCAATATATCTATGTTAGTTCAGAGACTTTCATTTGCGCCACAGTTCAGTGC 9840
Db 9781 AGGGCTCAATATATCTATGTTAGTTCAGAGACTTTCATTTGCGCCACAGTTCAGTGC 9840
Qy 9841 AGGCAATCTGGAATTTTCTCTCGCTGACAGATTTCCAGTCAATCCCACTTTGTACAGTTC 9900
Db 9841 AGGCAATCTGGAATTTTCTCTCGCTGACAGATTTCCAGTCAATCCCACTTTGTACAGTTC 9900

Qy 9901 TGGCACTTTTGGGTGAGGCGGTGATCCAAGAGCAGAAATTTCCAGCTATGTCAGGAG 9960
Db 9901 TGGCACTTTTGGGTGAGGCGGTGATCCAAGAGCAGAAATTTCCAGCTATGTCAGGAG 9960
Qy 9961 TGGCTGACCGTCCCACTGACACTCAACAAAGGCGAAACACACAGAGTGGCTTTTG 10020
Db 9961 TGGCTGACCGTCCCACTGACACTCAACAAAGGCGAAACACACAGAGTGGCTTTTG 10020
Qy 10021 TTGAATTTGAGTGTGGCCAGAGGGGCTGCACCTACTGATTTGACACAGGCAACA 10080
Db 10021 TTGAATTTGAGTGTGGCCAGAGGGGCTGCACCTACTGATTTGACACAGGCAACA 10080
Qy 10081 TTAATCTCAGCAAGTGAATTTGCAAGCCATTAATTTGAATTAATGATGATCAATGCA 10140
Db 10081 TTAATCTCAGCAAGTGAATTTGCAAGCCATTAATTTGAATTAATGATGATCAATGCA 10140
Qy 10141 ATCAGTATCAACAAGTGTGTTGGCTTTGGAAGATGAGTCTAGGGGCTTACAGAGTAC 10200
Db 10141 ATCAGTATCAACAAGTGTGTTGGCTTTGGAAGATGAGTCTAGGGGCTTACAGAGTAC 10200
Qy 10201 TACTCTATGAGTGTGATTTTGAAGCAGACACTGTGAAGAGCTGCTCTTAAGA 10260
Db 10201 TACTCTATGAGTGTGATTTTGAAGCAGACACTGTGAAGAGCTGCTCTCTTAAGA 10260
Qy 10261 GCCTGTAACATTTAGGTCATTTTTCAGTGCACCTTCTGAAAGTCTGACAGTCCCA 10320
Db 10261 GCCTGTAACATTTAGGTCATTTTTCAGTGCACCTTCTGAAAGTCTGACAGTCCCA 10320
Qy 10321 TGCAAAGCTGGCCAAACATGAGCACTTCCATGATTAATTAATTAATGAGGCTACTGCT 10380
Db 10321 TGCAAAGCTGGCCAAACATGAGCACTTCCATGATTAATTAATTAATGAGGCTACTGCT 10380
Qy 10381 TCTTGCAACACTGCTCTTCTCAAAATGAACTCAACAAACAAATTTCAAGTCTAGTAA 10440
Db 10381 TCTTGCAACACTGCTCTTCTCAAAATGAACTCAACAAACAAATTTCAAGTCTAGTAA 10440
Qy 10441 GTTAACAGTTTGAATGTCAATTAATAAAGTATATCTGCTTTCAGTGTGACGTTAT 10500
Db 10441 GTTAACAGTTTGAATGTCAATTAATAAAGTATATCTGCTTTCAGTGTGACGTTAT 10500
Qy 10501 GCCCACTGAACATCTGTATACAGCTGAAACCTGAGGCTCCAGATTTGTTAAACCT 10560
Db 10501 GCCCACTGAACATCTGTATACAGCTGAAACCTGAGGCTCCAGATTTGTTAAACCT 10560
Qy 10561 ACTTTATCAATCATAGATCATAGAAATGAGTGGGTTGGAAGGAGCCCAAGATCATG 10620
Db 10561 ACTTTATCAATCATAGATCATAGAAATGAGTGGGTTGGAAGGAGCCCAAGATCATG 10620
Qy 10621 AAGATCCAAACCCCGCCAGCAGGAGGCCAACCTCCAGATCTGTTACTAGACAG 10680
Db 10621 AAGATCCAAACCCCGCCAGCAGGAGGCCAACCTCCAGATCTGTTACTAGACAG 10680
Qy 10681 GCAGGCCAAGGGCTCATCAACCTGAGCCATGAAACCTCCAGAGATGAGACATCCACAA 10740
Db 10681 GCAGGCCAAGGGCTCATCAACCTGAGCCATGAAACCTCCAGAGATGAGACATCCACAA 10740
Qy 10741 CTCTGTGGGCAAGCTGTGCGACACCTCAGCACCCTCTCTGTAAGACTTTTCCCTGAC 10800
Db 10741 CTCTGTGGGCAAGCTGTGCGACACCTCAGCACCCTCTCTGTAAGACTTTTCCCTGAC 10800
Qy 10801 ATCCATCTAAGCCTTCCCTCTTGAGGTTAGATCCACTCCCTTGTGCTATCAGTGC 10860
Db 10801 ATCCATCTAAGCCTTCCCTCTTGAGGTTAGATCCACTCCCTTGTGCTATCAGTGC 10860
Qy 10861 TACTCTGTAAGAAATGTAATTTCTCTCTTTTGAAGAGTTGCAATGAGTCTCTTGA 10920
Db 10861 TACTCTGTAAGAAATGTAATTTCTCTCTTTTGAAGAGTTGCAATGAGTCTCTTGA 10920
Qy 10921 GCTTCTTCTCTCTGCGAGATGAACAGCCAGCTCCCTCAGCTCTTCTTATAGAGAGA 10980
Db 10921 GCTTCTTCTCTCTGCGAGATGAACAGCCAGCTCCCTCAGCTCTTCTTATAGAGAGA 10980

```

Oy 10981 GGTCGTCAGCCCTTGATCATCTTTGTGGCCCTCCTGTGAGACCCGCTCCAGAGCTCCA 11040
    |||||||
Db 10981 GGTCGTCAGCCCTTGATCATCTTTGTGGCCCTCCTGTGAGACCCGCTCCAGAGCTCCA 11040
Oy 11041 CATCTTCTGTACTGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 11100
    |||||||
Db 11041 CATCTTCTGTACTGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 11100
Oy 11101 GAGCAGAGAGAAAGAGGAGCAATCACTTCTGACACCTGCTGGCAGACCCCTCTCTATG 11160
    |||||||
Db 11101 GAGCAGAGAGAAAGAGGAGCAATCACTTCTGACACCTGCTGGCAGACCCCTCTCTATG 11160
Oy 11161 AGCCCTGATACACTGCTTCTGTGAGTCAACTTCTCTTATCAGTTCCATATTA 11220
    |||||||
Db 11161 AGCCCTGATACACTGCTTCTGTGAGTCAACTTCTCTTATCAGTTCCATATTA 11220
Oy 11221 ACAGGACAAATCAACAGAGTGTGATGCCAGTGCAGAGCTTTTTCACACTTCTTCATTT 11280
    |||||||
Db 11221 ACAGGACAAATCAACAGAGTGTGATGCCAGTGCAGAGCTTTTTCACACTTCTTCATTT 11280
Oy 11281 GGTAGATCTTATAGTGAAGAGTGTGAGTGTGCTTCTGCTGCTGCTCTCTCTCA 11340
    |||||||
Db 11281 GGTAGATCTTATAGTGAAGAGTGTGAGTGTGCTTCTGCTGCTGCTCTCTCTCA 11340
Oy 11341 ATACTCTGCTGATATACCTACACCCACCTGCACTGAATGCTCATGAGCCCTGACAG 11400
    |||||||
Db 11341 ATACTCTGCTGATATACCTACACCCACCTGCACTGAATGCTCATGAGCCCTGACAG 11400
Oy 11401 CAGGGCCCTGATGAACCCGGCAGCTGCTTATGATGCTGTTTAAATAGCAGATGACCAAG 11460
    |||||||
Db 11401 CAGGGCCCTGATGAACCCGGCAGCTGCTTATGATGCTGTTTAAATAGCAGATGACCAAG 11460
Oy 11461 TTGCACCTATGATACACAAATGTTGATCTCTGACAGCTGAGAGAGAGGCA 11520
    |||||||
Db 11461 TTGCACCTATGATACACAAATGTTGATCTCTGACAGCTGAGAGAGAGGCA 11520
Oy 11521 AATTGTCATTTGTCAGAAATGTTAGTAATCTGCAATTAACCTGTTTATCTACCA 11580
    |||||||
Db 11521 AATTGTCATTTGTCAGAAATGTTAGTAATCTGCAATTAACCTGTTTATCTACCA 11580
Oy 11581 TGGCTGTTTTTATGGCTGTTAGTAGTGACACTGATGTAACAATGGCTATGACGTAA 11640
    |||||||
Db 11581 TGGCTGTTTTTATGGCTGTTAGTAGTGACACTGATGTAACAATGGCTATGACGTAA 11640
Oy 11641 AATCAAGACTAGATATTTGCAACAGACTATTAATCTCTGCTGGCTTAGCCAAATGTGG 11700
    |||||||
Db 11641 AATCAAGACTAGATATTTGCAACAGACTATTAATCTCTGCTGGCTTAGCCAAATGTGG 11700
Oy 11701 TACTTCCCATTTGTATAGAATAATTTGGCAAGTTTAGACCAATGTTGAAGTGTGGAA 11760
    |||||||
Db 11701 TACTTCCCATTTGTATAGAATAATTTGGCAAGTTTAGACCAATGTTGAAGTGTGGAA 11760
Oy 11761 ATTCTGTATCTCAAGAGGGGCTTTTACAACTGTACAGAGAGGAATCAAAAGGGGG 11820
    |||||||
Db 11761 ATTCTGTATCTCAAGAGGGGCTTTTACAACTGTACAGAGAGGAATCAAAAGGGGG 11820
Oy 11821 TGGGAGAGAAATTAAGAGAGGAGGAGTGCAGAGAGCTTGCAGTCCCTGTGTAGC 11880
    |||||||
Db 11821 TGGGAGAGAAATTAAGAGAGGAGGAGTGCAGAGAGCTTGCAGTCCCTGTGTAGC 11880
Oy 11881 ACAGTGGCAACATGAGAGTCTTGTCTAATCTTGTGCTTGTCTGCTGCTGCTGCTGCT 11940
    |||||||
Db 11881 ACAGTGGCAACATGAGAGTCTTGTCTAATCTTGTGCTTGTCTGCTGCTGCTGCTGCT 11940
Oy 11941 TAGGG 11945
    |||||||
Db 11941 TAGGG 11945

```

```

RESULT 3
AF410481 18343 bp DNA linear VRF 30-JAN-2002
LOCUS AF410481
DEFINITION Gallus gallus egg white lysozyme and GAS41 (GAS41) genes,

```

```

regulatory regions and complete cds.
AF410481
AF410481.1 GI:18419435
VERSION
KEYWORDS
SOURCE
ORGANISM
    Gallus gallus.
    Gallus gallus.
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Phasianidae; Aves; Neognathae; Galliformes; Phasianidae;
    Phasianinae; Gallus.
    1 (bases 1 to 18343)
    Chong, S., Riggs, A.D. and Bonifer, C.
    The chicken lysozyme chromatin domain contains a second, widely
    expressed gene
    Nucleic Acids Res. 30 (2), 463-467 (2002)
JOURNAL
MEDLINE
PUBMED
    11788708
    2 (bases 1 to 18343)
    Bonifer, C. and Stippel, A.E.
    Direct Submission
    Submitted (13-AUG-2001) Molecular Medicine Unit, University of
    Leeds, Beckett Street, Leeds LS9 7TF, UK
FEATURES
    source
        1..18343
        /organism="Gallus gallus"
        /db_xref="taxon:9031"
        /chromosome="1"
        join(8985..9149,10387..10548,12343..12421,12503..12681)
        /product="egg white lysozyme"
        join(9014..9149,10387..10548,12343..12421,12503..12569)
        /note="clys"
        /product="egg white lysozyme"
        /protein_id="AA169327.1"
        /db_xref="GI:18419437"
        /translation="MRSLLIVLCEPLAALGKVFGRCELAAMKRHGLDNYRGYSLG
        NNVCAKESFNQATNRNTDGSFDYGLQINSRWCMNDGRTPSRNLGNIPCSALL
        SSDIATVNCARKIVSDGNMAMVAMRNRCRGTDVQAWIRCRLL"
        12891..15660
        /gene="GAS41"
        join(12891..12965,13108..13227,13324..13390,13468..13562,
        14156..14248,14332..14419,15096..15660)
        /gene="GAS41"
        /product="GAS41"
        join(12915..12965,13108..13227,13324..13390,13468..13562,
        14156..14248,14332..14419,15096..15660)
        /gene="GAS41"
        /note="nuclear protein; similar to glioma-amplified
        sequence 41; cgas41"
        sequence 41; cgas41"
        /codon_start=1
        /product="GAS41"
        /protein_id="AA169326.1"
        /db_xref="GI:18419436"
        /translation="MFKRMAEFGPDGSGHVGKVTIVKPYVGNVARYFGKREDEGHT
        HMTVYVKPYRNMEDMSAYVKIQFQLHESYGMPLRVYTPPEIETEGGEFEIIRI
        FLIDPNEREVLTILKLFQSDTNILGKTYVSEYDGMFOODPTAMMOQLTTSRO
        LFLGKHETFEPADELVKTRKRLPAKKTSSEIAMELRKLAASETINCLKSEIRKL
        EEDDSKDM"
BASE COUNT
    4685 a 4173 c 4345 g 5107 t 33 others
ORIGIN
Query Match 66.8%; Score 7974.4; DB 5; Length 18343;
Best local Similarity: 95.4%; Pred. No. 0;
Matches 8709; Conservative 2; Mismatches 314; Indels 106; Gaps 46;
Oy 2857 TCTAGACTGCATAGCTTGTGATCTGAATGACGTCTGATTGGCATGAGACACAGCA 2916
    |||||||
Db 1 1 TCTAGACTGCATAGCTTGTGATCTGAATGACGTCTGATTGGCATGAGACACAGCA 60
Oy 2917 CTTCTCATCTTACTTAACCTCACTTTTGGAAAGAAAGTAAAGCAAGGCGACAGTCC 2976
    |||||||
Db 61 CTTCTCATCTTACTTAACCTCACTTTTGGAAAGAAAGTAAAGCAAGGCGACAGTCC 120
Oy 2977 ATGAAATAGACAGTGCCTCAGGAGAAAGTGAACCTGATTTCTTTGGCTAGTGTCT 3036

```

```
Db 121 ATGAATATGAGACAGTGGCTCAGAGAAAGTGAACCTGGATTCTTGGCTAGTCTCT 180
Qy 3037 AAATCTG--TAGTGAGGAAAGTAACACCAGATCTCTTGAAGGGCTCCAGCTTTAATGCT 3094
Db 181 AAATCTGGTAGAGGAAAGTAACACCAGATCTCTTGAAGGGCTCCAGCTTTAATGCT 240
Qy 3095 TCCAAATTTGAAGTGGCAGAGCACTTGGCCACTGGTTATTTACTGCATTTATGCTCCAGTT 3154
Db 241 TCCAAATTTGAAGTGGCAGAGCACTTGGCCACTGGTTATTTACTGCATTTATGCTCCAGTT 300
Qy 3155 TCGGAGCTAACCTGGCTTCTCCACTATTGAGCATGAGCATATGCGCTTGGCTTCCAGAGCCA 3214
Db 301 TCGGAGCTAACCTGGCTTCTCCACTATTGAGCATGAGCATATGCGCTTGGCTTCCAGAGCCA 360
Qy 3215 GGTGAAGTGGGATGGGTGGAAGAGTGGCTGGTGGCTGGGGGAGCTGTGGGAGCT 3274
Db 361 GGTGAAGTGGGATGGGTGGAAGAGTGGCTGGTGGCTGGGGGAGCTGTGGGAGCT 420
Qy 3275 CCAAGCTGAGCTTGGGGTGGGAGCAGAGGAAAGTGGGTAACATATTTTAAAGTACT 3334
Db 421 CCAAGCTGAGCTTGGGGTGGGAGCAGAGGAAAGTGGGTAACATATTTTAAAGTACT 480
Qy 3335 GTGTTGCAAGCTGTCATCTGCAATAGCTAGGCTGCTGACTCTGGAAGATTAACAGTGT 3394
Db 481 GTGTTGCAAGCTGTCATCTGCAATAGCTAGGCTGCTGACTCTGGAAGATTAACAGTGT 532
Qy 3395 GGGTTTCAGTAAATATATGAGATGAATTCACAGTGGAGCACTTCAAGGGTAGATCATCTAACG 3454
Db 533 GGGTTTCAGTAAATATATGAGATGAATTCACAGTGGAGCACTTCAAGGGTAGATCATCTAACG 592
Qy 3455 ACACCAAGATCATCAAGCTATATGTTGGAAGCGGTATCAGAGAGCCAGAGAGGTAAGCACT 3514
Db 593 ACACCAAGATCATCAAGCTATATGTTGGAAGCGGTATCAGAGAGCCAGAGAGGTAAGCACT 652
Qy 3515 CTTCATATGTTTCCCTCCACGTAAGAGCAGTGGGAAAGTGAACCCCTTAGAGAGAGA 3574
Db 653 CTTCATATGTTTCCCTCCACGTAAGAGCAGTGGGAAAGTGAACCCCTTAGAGAGAGA 712
Qy 3575 CAAGGAAATTAATTCAGAGCATGTCTAGAGAACTTCTTTCGTAATTTCTACTTGCAG 3634
Db 713 CAAGGAAATTAATTCAGAGCATGTCTAGAGAACTTCTTTCGTAATTTCTACTTGCAG 772
Qy 3635 AGCTTTGATGCCCTGGCTTCTGGTGGCTTCTGAGACACCTGCAAGGCCCAAGACCTGTGGT 3694
Db 773 AGCTTTGATGCCCTGGCTTCTGGTGGCTTCTGAGACACCTGCAAGGCCCAAGACCTGTGGT 832
Qy 3695 GAGCTGAGAGGAAAGTATCTGCTCAAGTCCAAAGCTTTCAGCAGGATTAATTTGCTGCTCT 3754
Db 833 GAGCTGAGAGGAAAGTATCTGCTCAAGTCCAAAGCTTTCAGCAGGATTAATTTGCTGCTCT 892
Qy 3755 TCCCCAGACACTGTGACAGAGAGTGAACATGTCGAAGCCCTCTGTCGCCACTACTGTT 3814
Db 893 TCCCCAGACACTGTGACAGAGTGAACATGTCGAAGCCCTCTGTCGCCACTACTGTT 952
Qy 3815 GCTGAGGACAGACTGCTCTCAGAAAAAGAGACTTACTCTATGCCATAGCTCTGAAGGTAA 3874
Db 953 GCTGAGGACAGACTGCTCTCAGAAAAAGAGACTTACTCTATGCCATAGCTCTGAAGGTAA 1012
Qy 3875 AATGGGTTTTAAAA--AAGAAAAACAAGGCAAAACGGGTGGCCCATGGAAGAAAG 3931
Db 1013 AATGGGTTTTAAAAAGAAAGAAACACAAAGGCANAAACGCTGTGACACTGGAAGAAAG 1072
Qy 3932 CAGTGGTAAACATGGTAGAAAAAGTGCAGAAAGCCCCAGGAGTGTGACAGAGCCCTCTCT 3991
Db 1073 CAGTGGTAAACATGGTAGAAAAAGTGCAGAAAGCCCCAGGAGTGTGACAGAGCCCTCTCT 1132
Qy 3992 GCCACCTTAGAGGCGGGAACAAGCTTCCCTGCTAGGGCTGTGCCCCGGAAGTGGCTGTTT 4051
Db 1133 GCCACCTTAGAGGCGGGAACAAGCTTCTGCTTA--GGCTTGCCGCGCAAAAGTGGCTGTTT 1191
Qy 4052 CTTTGGTGGGTTTTGTTGGGTTTTGTTTTGAGATTTAGACACAAGGAAAGCTTGAAG 4111
Db 1191 CTTTGGTGGGTTTTGTTGGGTTTTGTTTTGAGATTTAGACACAAGGAAAGCTTGAAG 4111
Db 1192 CTTTGGTGGGTTTTGTTGGGCGGTGGTTTTGAGATTTAGACACAAGGAAAGCTTGAAG 1251
Qy 4112 GAGGTTTGGGACATATTTTGGTTTTGTAAAGCCTGACTTCAAAATATATTTTGGAG 4171
Db 1252 GAGGTTTGGGACATATTTTGGTTTTGTAAAGCCTGCTGCTTCAAAATATATTTTGGAG 1311
Qy 4172 GAGTGTAGCGAATTTGGCCAATTTAAATTAAGTTGCAAGAGATGGAAGCTGAGTAGTTG 4231
Db 1312 GAGTGTAGCGAATTTGGCCAATTTAAATTAAGTTGCAAGAGATGGAAGCTGAGTAGTTG 1371
Qy 4232 AGAGGTAACAGCTTTAATGAGATCTTTCGAAGCTCTGCTTCAAAAGCTGTTGAGT 4291
Db 1372 AGAGGTAACAGCTTTAATGAGATCTTTCGAAGCTCTGCTTCAAAAGCTGTTGAGT 1431
Qy 4292 GGTGAGACCTTGGATAGGATGAGTGTCTTGTATACATGTGTGATGCACTGTTGCTT 4351
Db 1432 GATGAGACCTTGGATAGGATGAGTGTCTTGTATACATGTGTGATGCACTGTTGCTT 1491
Qy 4352 TCCATCCACATCCATGATTCACATCCAGCATGTCAGCATTTGTACATTCATTCGTATTA 4411
Db 1492 TCCATCCACATCCATGATTCACATCCAGCATGTCAGCATTTGTACATTCATTCGTATTA 1551
Qy 4412 TCTGACATACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4471
Db 1552 TCTGACATACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1611
Qy 4472 GCCCAGATTTGAGAAGATGGCAGTGTGCTTCCCTTTCCTGCTGCTGCTGCTGCTGCTGCT 4531
Db 1612 GCCCAGATTTGAGAAGATGGCAGTGTGCTTCCCTTTCCTGCTGCTGCTGCTGCTGCTGCT 1667
Qy 4532 TCTCTGGCTTTGACACCTCAGCAAAATAGTCTTCTGCTTTCATTTCTGGCATTAATTC 4591
Db 1668 TCTCTGGCTTTGACACCTCAGCAAAATAGTCTTCTGCTTTCATTTCTGGCATTAATTC 1727
Qy 4592 AAATATCTTTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4651
Db 1728 AAATATCTTTGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1787
Qy 4652 CTTAGAGTGAAGAGAGAGAGAGAGATGTTGGCCGAGTTCCTGATGAACACACCTC 4711
Db 1788 CTTAGAGTGAAGAGAGAGAGAGAGATGTTGGCCGAGTTCCTGATGAACACACCTC 1847
Qy 4712 TGAATTA--TGGCCAAAGTGGGTGGTTTTCTGAGGAACGGGACGCTTTCCTCTGA 4770
Db 1848 TGAATTAATTTGGCCAAAGTGGGTGGTTTTCTGAGGAACGGGACGCTTTCCTCTGA 1907
Qy 4771 AGCAAGAGACTCTGGGAGTTCAGTATTTTGAACATGATGATGAGTGGGCTTGA 4830
Db 1908 AGCAAGAGACTCTGGGAGTTCAGTATTTTGAACATGATGATGAGTGGGCTTGA 1967
Qy 4831 GCAGATTCCTTAGTTCCTGCTACTTCTTCTTCTTGGCAGTCAAGTTATTTCTGAC 4890
Db 1968 GCAGATTCCTTAGTTCCTGCTACTTCTTCTTCTTGGCAGTCAAGTTATTTCTGAC 2027
Qy 4891 AGACAAACAGCCACCCCACTCAGGCTTGAAGATATGCTGCTGCTGGGTGTGA 4950
Db 2028 AGACAAACAGCCACCCCACTCAGGCTTGAAGATATGCTGCTGCTGGGTGTGA 2087
Qy 4951 CAGCTCTGCGCTTGGTGAAGGGGATTAAGGGGACCACTTCATCCCAAGAGATCCCTC 5010
Db 2088 CAGCTCTGCGCGTGAAGGGGATTAAGGGGACCACTTCATCCCAAGAGATCCCTC 2147
Qy 5011 ATTCAATGATCAAGCTGTAAAGAACTTGGGCTCCAAACCTCAAAACATTAATTTGAGTAG 5070
Db 2148 ATTCAATGATCAAGCTGTAAAGAACTTGGGCTCCAAACCTCAAAACATTAATTTGAGTAG 2207
Qy 5071 AATGTAATTAACATGCTATTCGCTATTCCTAAGTCAATTTAGTCTGAGCTGCAAGCATG 5130
Db 2208 AATGTAATTAACATGCTATTCGCTATTCCTAAGTCAATTTAGTCTGAGCTGCAAGCATG 2267
Qy 5131 TAGGTGCGAGCTCCACATTTCTCAAAAGCCACGATGAGAGAGATGAATAATGAGAGC 5190
Db 2268 TAGGTGCGAGCTCCACATTTCTCAAAAGCCACGATGAGAGAGATGAATAATGAGAGC 2327
```

| | | | | | |
|----|------|-----------------|--|----------------------------|------|
| QY | 5191 | CGATTTCAGAACAA | CCACAGGAGCTGTTGCCGAGAAAC | CTGATGAAATATGATGAATTG | 5250 |
| Db | 2328 | CGATTTCAGAACAA | CCACAGGAGCTGTTGCCGAGAAAC | CTGATGAAATATGATGAATTG | 2386 |
| QY | 5251 | TGTGGTGGACATTT | TTTTTAAATACATTAAC | CTCAATCTCAATGTAGAGTGCAGT | 5310 |
| Db | 2387 | TGTGGTGGACATTT | TTTTTAAATACATTAAC | CTCAATCTCAATGTAGAGTGCAGT | 2436 |
| QY | 5311 | GTTTTTAATTAGAC | AGCCATTAACCCAGGTGAGCGAGTAC | CAATTTTTTCTCTACAAAGAAAAACG | 5370 |
| Db | 2437 | GTTTTTAATTAGAC | AGCCATTAACCCAGGTGAGCGAGTAC | CAATTTTTTCTCTACAAAGAAAAACG | 2496 |
| QY | 5371 | ATTTCGAGCTCTGC | GTAACTATTAAGTTCTCATAGGGGCTGAAGCTC | CCCCCTGCGCTGCC | 5430 |
| Db | 2497 | ATTTCGAGCTCTGC | GTAACTATTAAGTTCTCATAGGGGCTGAAGCTC | CCCCCTGCGCTGCC | 2556 |
| QY | 5431 | TGSCATCTCAGCTG | G-----AGTGCAGTGCATTTCTTGGGTTCCTCTCACAGCAGTA | | 5485 |
| Db | 2557 | TGSCATCTCAGCTG | GAGTGAGTGCAGTGCATTTCTTGGGTTCCTCTCACAGCAGTA | | 2616 |
| QY | 5486 | ATGGGCAATFACTT | CACAAAA-----TTCCTTTCTTCTCTGTCAATGTGGATCCCTAC | | 5539 |
| Db | 2617 | ATGGGCAATFACTT | CACAAAAATCTTCTTCTTCTCTGTCAATGTGGATCCCTAC | | 2676 |
| QY | 5540 | TGTGCCCTCTGGTT | TATAGTTACCCCTGACCTGTTCCATTCACAGCGTTTGGAAGAGAA | | 5599 |
| Db | 2677 | TGTGCCCTCTGGTT | TATAGTTACCCCTGACCTGTTCCATTCACAGCGTTTGGAAGAGAA | | 2736 |
| QY | 5600 | AAAGAAATTTGGAAT | TAACACATGTCTACGTTATCACCCTCTCCAGCATTTTGTTTTTAA | | 5659 |
| Db | 2737 | AAAGAAATTTGGAAT | TAACACATGTCTACGTTATCACCCTCTCCAGCATTTTGTTTTTAA | | 2796 |
| QY | 5660 | TATATGCATTAATCT | GGCTTAAGATTTGGAATGAGGGGGTGGTATTTACCGAGAA | | 5719 |
| Db | 2797 | TATATGCATTAATCT | GGCTTAAGATTTGGAATGAGGGGGTGGTATTTACCGAGAA | | 2856 |
| QY | 5720 | CAAGGAGGCTTATAT | ATAAATCACTCACTTTTATTTAGAGACTGSCAGCTGTCAAAA | | 5779 |
| Db | 2857 | CAAGGAGGCTTATAT | ATAAATCACTCACTTTTATTTAGAGACTGSCAGCTGTCAAAA | | 2916 |
| QY | 5780 | CAAAAAAGGCTTAC | CCACAAATTAAGGATPACCGCTATAGCAGGAGAGGCCACGACGA | | 5839 |
| Db | 2917 | CAAAAAAGGCTTAC | CCACAAATTAAGGATPACCGCTATAGCAGGAGAGGCCACGACGA | | 2976 |
| QY | 5840 | GGGATGTCACACT | GTGGCACTATGCGACGCGCTGCTTGTGACTCTGAGACCACTGCTT | | 5899 |
| Db | 2977 | GGGATGTCACACT | GTGGCACTATGCGACGCGCTGCTTGTGACTCTGAGACCACTGCTT | | 3036 |
| QY | 5900 | TGGAAATGACAGCA | CTTGGTGGCAATTTCCCTTGTTCAGAAATGCGTATGAGCGTGTGCTTG | | 5959 |
| Db | 3037 | TGGAAATGACAGCA | CTTGGTGGCAATTTCCCTTGTTCAGAAATGCGTATGAGCGTGTGCTTG | | 3096 |
| QY | 5960 | GGCAGAGTTTCTT | TAATTAAGGCACATCTTTTCCCTCTCTCCATTCCTCAACAT | | 6019 |
| Db | 3097 | GGCAGAGTTTCTT | TAATTAAGGCACATCTTTTCCCTCTCTCCATTCCTCAACAT | | 3156 |
| QY | 6020 | GTCTCCATGCTGTA | TATCCAGTCAAGTGAAGCTTCAACAATGAATTCATCACTGTAGG | | 6079 |
| Db | 3157 | GTCTCCATGCTGTA | TATCCAGTCAAGTGAAGCTTCAACAATGAATTCATCACTGTAGG | | 3216 |
| QY | 6080 | ATTCTCGTGTGAT | CAAAATCTTGTGTGAGGCTGATTAATA--TATGGAAGCTTATTAAT | | 6137 |
| Db | 3217 | ATTCTCGTGTGAT | CAAAATCTTGTGTGAGGCTGATTAATAATATGTAAGCTTATTAAT | | 3276 |
| QY | 6138 | TTTTCGTTCTTCAT | PATACAGTCTCTGTATACAAATTCACATTCACACACAG--CAATTTAA | | 6196 |
| Db | 3277 | TTTTCGTTCTTCAT | PATACAGTCTCTGTATACAAATTCACATTCACACACAGCAATTTAA | | 3336 |
| QY | 6197 | GGTGAAGAGGCTG | TGGTGGATGAAGAGGCTTCTTAGCTTTCTTCTTGCAAGGC | | 6256 |
| Db | 3337 | GGTGAAGAGGCTG | TGGTGGATGAAGAGGCTTCTTAGCTTTCTTCTTGCAAGGC | | 3396 |

| | | | |
|----|------|--|------|
| QY | 6257 | CACAGGAAATGCTGAGAGCTGTAGATACATACGCTGGGTAAAGAAATTGACTCTCTGCT | 6316 |
| Db | 3397 | CACAGGAAATGCTGAGAGCTGTAGATACAGGCT - GGGTAAAGAAATTGACTCTCTGCT | 3455 |
| QY | 6317 | GGGAGAGTAAAGCCATCTTATACCCCTTCAGACATCTTAAAGCAAAATAGGCTC | 6376 |
| QY | 6377 | TATCTGGGGTTTTGTCTCCGCTGTTCCCTCCCTGGAAGGCTATCTACATTTTCACTGCTC | 6436 |
| Db | 3515 | TATCTGGGGTTTTGTCTCCGCTGTTCCCTCCCTGGAAGGCTATCTTAAAGCTTAA - GGGTTC | 3514 |
| QY | 6437 | CCACGGTTACAACCAAGATACAGCCGTGAATTTTTTCTAGGCACTTACATAAATTTG | 6496 |
| Db | 3575 | TGTCCGCTGGAAGACTGTGTATTACTTAAGTGTCCACACTTCGGTCTCTGCGCGCTGG | 3634 |
| QY | 6497 | ACCTGGTA - CCAATATGTCTCTATATATATTATTTTCTCCCTGCTGTTTAAACCTC | 6555 |
| Db | 3635 | CGTTTCCAGGTACCATTTGTCTTATATATATATATTTTCTCCCTGCTGTTTAAACCTC - | 3693 |
| QY | 6556 | TAAAGCATTTGAGAACACATAGAAATCATAGATGTTTGGATTGGAAGGGGCTTTAAACAT | 6615 |
| Db | 3694 | TAAAGCATTTGAGAACACATAGAAATCATAGATGTTTGGATTGGAAGGGGCTTTAAAT | 3753 |
| QY | 6616 | CATCCATTTCCAAACCTCTGSCCATAGGGCTTTGCCACCACTGAGCTCAGCTGCCAGG | 6675 |
| Db | 3754 | CATCCATTTCCAAACCTCTGSCCATAGGGCTTTGCCACCACTGAGCTCAGCTGCCAGG | 3813 |
| QY | 6676 | GCCCATTCACCCCTGGGCTTGAGACCTCCAGGGAATGGGGCAACCAAGCTTCCTGGGC | 6735 |
| Db | 3814 | GCCCATTCACCCCTGGGCTTGAGACCTCCAGGGAATGGGGCAACCAAGCTTCCTGGGC | 3873 |
| QY | 6736 | AGCCGTGCGCAACACCTCACCACATCTCTGGGTAAAGAAATTCCTTTTAACATCTAATCA | 6795 |
| Db | 3874 | AGCCGTGCGCAACACCTCACCACATCTCTGGGTAAAGAAATTCCTTTTAACATCTAATCA | 3928 |
| QY | 6796 | AATCTCTCTCTTTAGTTTAAAGCAATTCCTCTTTTCCGCTGCTATCTGCCAAGA | 6855 |
| Db | 3929 | GATCTCTCTCTTTAGTTTAAAGCAATTCCTCTTTTCCGCTGCTATCTGCCAAGA | 3988 |
| QY | 6856 | ATGTGTATTTG - TCTCCCTCTGCTTATAGCAGGAAGTACGGAAGCTGCAAGAGCT | 6914 |
| Db | 3989 | ATGTGTATTTGCTCCCTCTGCTTATAGCAGGAAGTACGGAAGCTGCAAGAGCT | 4048 |
| QY | 6915 | CTCCCAAGAGCTCTCTCTTCCAGGCTGAACAGCCACAGCTCTTCAAGCTCTCTTG - | 6973 |
| Db | 4049 | CTCCCAAGAGCTCTCTCTTCCAGGCTGAACAGCCACAGCTCTTCAAGCTCTCTTG | 4108 |
| QY | 6974 | TAGAGATCATTTAGTGCCCTCTCTGAGACCCATTCACAAGTTCCAGGGCTTTCTG | 7033 |
| Db | 4109 | TAGAGATCATTTAGTGCCCTCTCTGAGACCCATTCACAAGTTCCAGGGCTTTCTG | 4168 |
| QY | 7034 | TGGACCCCAAGCTGTGATGAGTCACTTCAGATGGGGCTTACAAAGCAGAGCAGATGG | 7093 |
| Db | 4169 | TGGACCCCAAGCTGTGATGAGTCACTTCAGATGGGGCTTACAAAGCAGAGCAGATGG | 4228 |
| QY | 7094 | GGACATGCTTTACCCCTCCCTGCTGCTCCCTGTTTGAATGAGCCACAGACTGT | 7153 |
| Db | 4229 | GGACATGCTTTACCCCTCCCTGCTGCTCCCTGTTTGAATGAG - CCAAGGACTGT | 4287 |
| QY | 7154 | TGGCCTTTCAAGGCTCCAGAACCCCTTGCTGATTTGTCAAGCTTTTCACTCCACAGAAC | 7213 |
| Db | 4288 | T - GCCTTTCAAGGCTCCAGAACCCCTTGCTGATTTGTCAAGCTTTTCACTCCACAGAAC | 4340 |
| QY | 7214 | CCAGGCTCTTGTTAACTTTCTGCCCTC - ACTCTGTAGCTGTTTCAGAGACTTC | 7272 |
| Db | 4341 | C-----TCTGTATTAATACTTCTGCCCTCAACTCTGTAACTGTTTCAAGAACTTC | 4394 |
| QY | 7273 | CATCTTTAGAGACAGACTGTGTTAACACTACACTACGCGCTATCTGTGATATATACATTTCA | 7332 |
| Db | 4395 | CATCTTTAGAGACAGACTGTGTTAACACTACACTACGCGCTATCTGTGATATATACATTTCA | 4454 |
| QY | 7333 | GTCTATGTTTCTCTAACAGGACGAAATATGTATTCCTCTAACAAATAATACATGAGAT | 7392 |

|||||
Db 4455 GTTCATGTTCTCGTAACAGGAGCAAGATATGATTCCTCAACAAATAATATGAGAAAT 4514
QY 7393 TCCATAGGCATCTCAGTAGGGTTTTTCATGGCAGTATTGACATATGCAATATTTGGTGA 7452
Db 4515 TCCATAGGCATCTCAGTAGGGTTTTTCATGGCAGTATTGACATATGCAATATTTGGTGA 4574
QY 7453 AGTACCTTCCAAAGCTGCGGCCCTCCCAATAATCTGTATTTGGGATCAGTTACCTTTGGG 7512
Db 4575 AGTACCTTCCAAAGCTGCGGCCCTCCCAATAATCTGTATTTGGGATCAGTTACCTTTGGG 4634
QY 7513 GTAACCTTTGATCTGAGAGAACCTTGGGGTTCTGATGTCTCTCAGCTCTGTCTGT 7572
Db 4635 GTAACCTTTGATCTGAGAGAACCTTGGGGTTCTGATGTCTCTCAGCTCTGTCTGT 4694
QY 7573 CTGACTGACCATTTTCTAGATCACCAAGTTGTCTGTACCAACTTCCCTGCTCCATC 7632
Db 4695 CTGACTGACCATTTTCTAGATCACCAAGTTGTCTGTACCAACTTCCCTGCTCCATC 4754
QY 7633 CTTCCTCCAGCTTGTATCTTTGACAAATACAGGCTATTTTGTGTGCTTCAGACCA 7692
Db 4755 CTTCCTCCAGCTTGTATCTTTGACAAATACAGGCTATTTTGTGTGCTTCAGACCA 4814
QY 7693 TTTAATTTCTGAGTGCATCTGTCTGTGATGCGACCTGGAACAGGATTTTCAGACGTC 7752
Db 4815 TTTAATTTCTGAGTGCATCTGTCTGTGATGCGACCTGGAACAGGATTTTCAGACGTC 4874
QY 7753 TTGCAAAAGACATCTAGCTGAAACCTTTCTGCCATTCATATTTCTACAGTTCTTCTG 7812
Db 4875 TTGCAAAAGACATCTAGCTGAAACCTTTCTGCCATTCATATTTCTACAGTTCTTCTG 4934
QY 7813 TTTGAGGAGGACCATTAATTTACTAGAACTTCTCAGTCTGACAACTTTATGCAATTTATAC 7872
Db 4935 TTTGAGGAGGACCATTAATTTACTAGAACTTCTCAGTCTGACAACTTTATGCAATTTATAC 4994
QY 7873 TTTCTATTATGACTACTTTTGACATACACAGACGCAATTTTGGTGGAATTTCCA 7932
Db 4995 TTTCTATTATGACTACTTTTGACATACACAGACGCAATTTTGGTGGAATTTCCA 5054
QY 7933 CAGTGTCTCTGTGTCTTTCACATGGTTTTTACTGTATCTTCCGTTATTAACCTTGGCAAT 7992
Db 5055 CAGTGTCTCTGTGTCTTTCACATGGTTTTTACTGTATCTTCCGTTATTAACCTTGGCAAT 5114
QY 7993 CTGCCACCTGCCCCATCACAAAGAAAGATTTCTTTTATTTACTTCTCTTACGCCAAT 8052
Db 5115 CTGCCACCTGCCCCATCACAAAGAAAGATTTCTTTTATTTACTTCTCTTACGCCAAT 5174
QY 8053 AAACAATAATGTAGAAAGCCCAAAAGAACTTGTGGGACAGGCTCCCATCAAGGAGAGA 8112
Db 5175 AAACAATAATGTAGAAAGCCCAAAAGAACTTGTGGGACAGGCTCCCATCAAGGAGAGA 5234
QY 8113 CAGCTGAAGGTTGTGATCTCAATAGAAATTAATAAGCTGTGTGACAGAGTT 8172
Db 5235 CAGCTGAAGGTTGTGATCTCAATAGAAATTAATAAGCTGTGTGACAGAGTT 5294
QY 8173 TTGGCTGATTTATACAGGACGCCCCAAGCCAGAGAGGCTGTCTCCCAAGGACCTTGGC 8232
Db 5295 TTGGCTGATTTATACAGGACGCCCCAAGCCAGAGAGGCTGTCTCCCAAGGACCTTGGC 5354
QY 8233 AGTCCCTGGTTTGTAAAGTAAAGTCAATAGTAACTTTCTGGGAATTTGCGTGTGAATCA 8292
Db 5355 AGTCCCTGGTTTGTAAAGTAAAGTCAATAGTAACTTTCTGGGAATTTGCGTGTGAATCA 5414
QY 8293 TGATGCGAGTTCTGTGCTTTTACTATGATGTAAGATGCTAAATATAGAGACAGCAAAAGTAAAC 8352
Db 5415 TGATGCGAGTTCTGTGCTTTTACTATGATGTAAGATGCTAAATATAGAGACAGCAAAAGTAAAC 5474
QY 8353 ACTGTCTCTGTAGTGTCTGTCTCTATCCAGACAGGATGGCAGCTGCAACCAAGATGAG 8412
Db 5475 ACTGTCTCTGTAGTGTCTGTCTCTATCCAGACAGGATGGCAGCTGCAACCAAGATGAG 5534
QY 8413 GGATGCTCCAGCTGAGGATGCTGGGGCAGTAAACAGTGGGCTCCATGCTGCTGCTCAT 8472
|||||

Db 5535 GGATGCTCCAGCTGAGGATGCTGGGGCAGTAAACAGTGGGTCCATGCTGCTGCTCAT 5594
QY 8473 TAGCATCACCTGAGCCCTGACCAGCCCATCAGAAGATATCCAAAGCTGAGAAAGTTG 8532
Db 5595 TAGCATCACCTGAGCCCTGACCAGCCCATCAGAAGATATCCAAAGCTGAGAAAGTTG 5654
QY 8533 CTCATCTTCTTGACATCATCAAAACCTTTGGCCCTGATGTGATGCTCCCGATGCTTAAATG 8592
Db 5655 CTCATCTTCTTGACATCATCAAAACCTTTGGCCCTGATGTGATGCTCCCGATGCTTAAATG 5714
QY 8593 TGGTCACTACATCTTTATTTTCTATGATTTCAAGTCAAGACCTCCGATCAGAGAGA 8652
Db 5715 TGGTCACTACATCTTTATTTTCTATGATTTCAAGTCAAGACCTCCGATCAGAGAGA 5774
QY 8653 ACACATAGTGGGAATGTACCTCAGCTCCAAAGCCGAGATCTTCTCAATGATCATGAT 8712
Db 5775 ACACATAGTGGGAATGTACCTCAGCTCCAAAGCCGAGATCTTCTCAATGATCATGAT 5834
QY 8713 GCTACTTAGGAAGTGTGTGTGTGATGTAGAAATGTGCTTTGTTATTTTCTTCTG 8772
Db 5835 GCTACTTAGGAAGTGTGTGTGTGATGTAGAAATGTGCTTTGTTATTTTCTTCTG 5894
QY 8773 CTGTGAGGAACATTTTGAATACAGAGAAAGAAAGTGCCTTCTGCAATGGGAGA 8832
Db 5895 CTGTGAGGAACATTTTGAATACAGAGAAAGAAAGTGCCTTCTGCAATGGGAGA 5954
QY 8833 GTTGTACACTTGCAAAATTAAGAGATGACGTCCCAATGTTGATATCTCAGAGTCTGAA 8892
Db 5955 GTTGTACACTTGCAAAATTAAGAGATGACGTCCCAATGTTGATATCTCAGAGTCTGAA 6014
QY 8893 GGAGATTCAGAAACGTGTATACAAATTTTACAGCTTCTGAAATGCAAGCTTTGAAAGCTG 8952
Db 6015 GGAGATTCAGAAACGTGTATACAAATTTTACAGCTTCTGAAATGCAAGCTTTGAAAGCTG 6074
QY 8953 TTTCGTGCGCGAGGACGTACTAGTACGAACCCCTCGGAAACAGGAACAAATGCTTCAAGT 9012
Db 6075 TTTCGTGCGCGAGGACGTACTAGTACGAACCCCTCGGAAACAGGAACAAATGCTTCAAGT 6134
QY 9013 GCAGCAGAGAGAAACCTTGGCCCATCATGAAGTAATTAACCTGCGCTGAAGAAAT 9072
Db 6135 GCAGCAGAGAGAAACCTTGGCCCATCATG - AAGTGAATTAACCTGCGCTGAAGAAAT 6193
QY 9073 CCAGCTCCTGTTTGAAGCAGGTGTGACACCTCCACACCTGGAACAACAGTTCATTTTAT 9132
Db 6194 CCAGCTCCTGTTTGAAGCAGGTGTGACACCTCCACACCTGGAACAACAGTTCATTTTAT 6253
QY 9133 AGGACTTCAGAGAAAGATCTTCTCTTAAGCTTTCTTAATTTATGATACATCTCAGTTGGC 9192
Db 6254 AGGACTT - CAGGAAGATCTTCTCTTAAGCTTTCTTAATTTATGATACATCTCAGTTGGC 6312
QY 9193 AGATGACTATGACTACTGACAGGAGAAATGAGAACTAGCTGGGAATTTTCTGTTTGACC 9252
Db 6313 AGATGACTATGACTACTGACAGGAGAAATGAGAACTAGCTGGGAATTTTCTGTTTGACC 6372
QY 9253 ACCATGAGATCACCATTCTTACTGTGATTTGGAATTAATTTTGAATTTGCAAGC 9312
Db 6373 ACCATGAGATCACCATTCTTACTGTGATTTGGAATTAATTTTGAATTTGCAAGC 6432
QY 9313 AAGAGTTAGCGAAGATCTTCAATTTCTCCATGTTGGTGACACAGAGTCTGCTATGAA 9372
Db 6433 AAGAGTTAGCGAAGATCTTCAATTTCTCCATGTTGGTGACACAGAGTCTGCTATGAA 6492
QY 9373 ACTGTGCTTACAGGAAGAGATAAATCATAGGATTAATTAATCTAAGTTTGAAGCA 9432
Db 6493 ACTGTGCTTACAGGAAGAGATTAATAATCATAGGATTAATTAATCTAAGTTTGAAGCA 6552
QY 9433 ATGAGGTTTATGCTGATTTGACATGAAGAAATTTAGACCTTACTGGATAGCTATGGA 9492
Db 6553 ATGAGGTTTATGCTGATTTGACATGAAGAAATTTAGACCTTACTGGATAGCTATGGA 6612
QY 9493 TTTAGGTTGCTTTTGTGTTAGTTACTTATTTGACCCAGCTGAGGTCAG - TTGGAATCA 9551
Db 6613 TTTAGGTTGCTTTTGTGTTAGTTACTTATTTGACCCAGCTGAGGTCAGTTAGGAATCA 6672
|||||

| | | | |
|----|-------|--|-------|
| QY | 9552 | GGCTCTCGGGCTACTGGCATGGATTGATTACATACACTGTAATTTTACAGCTGATTTA | 9611 |
| Db | 6673 | GGCTCTCAGGGCTACTGGGATGGATTGATTACATACACTGTAATTTTACAGCTGATTTA | 6732 |
| QY | 9612 | GGTTTATAGTACTTTTGACAGTAATTCATAGGGTAGTAATGTTATATCTCA-GGGAAA | 9670 |
| Db | 6733 | GGGTTTATAGTACTTTTGTAGCAAAATCATAGGGTTTAGTAATGTTATATCTCAGGGGAAA | 6792 |
| QY | 9671 | AAAAAAAAAGCCAAACCCTGACAGACATCCAGCTCAGGTGGAATCAAGGATCACAGCTC | 9730 |
| Db | 6793 | AAAAAAAAAGCCAAACCCTGACAGACATCCAGCTCAGGTGGAATCAAGGATCACAGCTC | 6852 |
| QY | 9731 | AGTGGGCTCCAGAGAACACAGGACACTCTCTCTTAGACACTTATATGTACAGGGCTCAA | 9790 |
| Db | 6853 | AGTGGGCTCCAGAGAACACAGGACACTCTCTCTTAGACACTTATATGTACAGGGCTCAA | 6912 |
| QY | 9791 | GATTAATGATGTTATATACAAAGACTTCCATTCTGGCCACAGTTCAGCTAGGCAATCTCT | 9850 |
| Db | 6913 | GGTAATGATGTTATATAGTACAGAGACTTTCATTCTGGCCACAGTTCAGGCAATCTCT | 6972 |
| QY | 9851 | GGATTTTCTCTCCGCTGCACAGTTCAGTCAATCCAGTTTGTACAGTTCTGCACATTTT | 9910 |
| Db | 6973 | GGATTTTCTCTCCGCTGCACAGTTCAGTCAATCCAGTTTGTACAGTTCTGCACATTTT | 7032 |
| QY | 9911 | TGGGTCAGGCGCGTATCCAAAGGAGCAAGATTCACACTATAGTCAGAGGATGCTGACCG | 9970 |
| Db | 7033 | TGGGTCAGGCGCGTATCCAAAGGAGCAAGATTCACACTATAGTCAGAGGATGCTGACCG | 7092 |
| QY | 9971 | TCCCAACTCAGTGCACACTCAAAAGGCGAAACACAGAGTGCGCTTTTGTGTAATTTGC | 10031 |
| Db | 7093 | TCCCAACTCAGTGCACACTCAAAAGGCGAAACACAGAGTGCGCTTTTGTGTAATTTGC | 7152 |
| QY | 10031 | AGTGGGCGCCAGAGGAGGCGACACCACTAGTGATTCACACAGGAGCAACCTTATCTCTCA | 10091 |
| Db | 7153 | AGTGGGCGCCAGAGGAGGCGACACCACTAGTGATTCACACAGGAGCAACCTTATCTCTCA | 7212 |
| QY | 10091 | GCAAGTGCATTTTGCAGCCATTTAAATTTGAATCACTCAATGCAATGCAATGCT-AT | 10141 |
| Db | 7213 | GCAAGTGCATTTTGCAGC--GTAATTTAACTGATTGACATGATGTTACGTTAGTCATTA | 7270 |
| QY | 10141 | CACAAGTGGTTTGGCTTGGAGATGCGATCTAGGGGCTCTACAGAGTATGCTACTCTCT | 10201 |
| Db | 7271 | GACAAGTGGTTTGGCTTGGAGATGCGATCTAGGGGCTCTACAGAGTATGCTACTCTCT | 7330 |
| QY | 10201 | AATGAGTGTGCAATTTTGAGCAGAGCACTGTGAAAAGCTGGCCCTGTAAGAAGCTGCTA | 10261 |
| Db | 7331 | AATGAGTGTGCAATTTTGAGCAGAGCACTGTGAAAAGCTGGCCCTGTAAGAAGCTGCTA | 7390 |
| QY | 10261 | AACATTTAGGTCAAATTTTCCAGTGCATTTTGTGAAGTGTGTGCAATTTCCCATGCAAAAGC | 10321 |
| Db | 7391 | AACATTTAGGTCAAATTTTCCAGTGCATTTTGTGAAGTGTGTGCAATTTCCCATGCAAAAGC | 7450 |
| QY | 10321 | TGCCCCAATATGCACTTCCATTTGAAATATCAATTTATATGCAAGCGTACTGCTTCTTGCCA | 10381 |
| Db | 7451 | TGCCCCAATATGCACTTCCATTTGAAATATCAATTTATATGCAAGCGTACTGCTTCTTGCCA | 7510 |
| QY | 10381 | GCA-CTGTCTCTCTCAAAATGAACCTCAACAACATTTTCAAAGTCTAGTAGAAAGTAACAA | 10441 |
| Db | 7511 | GCACGTCCTCTCTCAAAATGAACCTCAACAACATTTTCAAAGTCTAGTAGAAAGTAACAA | 7570 |
| QY | 10441 | GCTTTGATATGTCATT-AAAAAGTATATCTGCTTTCAGTAGTTCAGCTTATTTATG-CCCA | 10501 |
| Db | 7571 | GCTTTGATATGTCATTAAAAAGTATATCTGCTTTCAGTAGTTCAGCTTATTTATGCCCCA | 7630 |
| QY | 10501 | CTAGAAACATCTTGTGTCAGAGCTGAACACTGGGGCTCCAGATTGTGTGTAACCACTACTCT | 10561 |
| Db | 7631 | CTAGAAACATCTTGTGTCAGAGCTGAACACTGGGGCTCCAGATTGTGTGTAACCACTACTTA | 7690 |
| QY | 10561 | ATACATATCATAGATCATAGAATGGCTT-GGCTTGGAGAGGACCCCAAGATCATGAGA | 10621 |
| Db | 7691 | TACAAATATCATAGATCATAGAATGGCTTGGCTTGGAGAGGACCCCAAGATCATGAGA | 7750 |

[illegible]

| | | | |
|----|-----|--|-----|
| QY | 1 | TGCGCCCTTGGATGATTTACACCTGTGTTGATTTATTCATCTCTCTCTGGCCAAAGAAAGATA | 60 |
| | 122 | TGCGCCCTTGTGATGATTTACACCTGTGTTGATTTATTCATCTCTCTGTGCGCAAGAAAGAA-- | 179 |
| Db | | | |
| QY | 61 | TACACGTCGTATACACGCTCTGAGAAATACTTGSTATTCTCTGATCAGTGTTTT | 120 |
| | 180 | -----TATACACGCTCTGAGAAATACTTGSTATTCTCTGATCAGGTTTT | 228 |
| Db | | | |

| | | | |
|----|------|--|------|
| QY | 121 | ATGGTATGTTGAATATATGGATTAAGGCTGTCCTTTGCTTGGGAGACAACCCCA | 180 |
| Db | 230 | ATAAGTATATGTGAATATATGGATTAAGGCTGTCCTTTGCTTGGGAGACAACCCCA | 289 |
| QY | 181 | CAGCAGAGTGTGGTGGGGTGGTGGCGACGTCAGTGACAGAGAGGTTTTTTTGGCCGTTT | 240 |
| Db | 280 | CAGCAGGTGGTGGTGGGGTGGTGGCGACGTCAGTGACAGAGAGGTTTTTTTGGCCGTTT | 348 |
| QY | 241 | TTTTTT-TTTTTTTTTTTTTTTTTAAGTAAGTGTCCTTTTTCTTAAATATTTTCTACTGG | 299 |
| Db | 349 | TTTTTTGTTTTTTTTTTTTTTTAAGTAAGTGTCCTTTTTCTTAAATATTTTCTACTGG | 408 |
| QY | 300 | ACTGATGTTTTGACAGTCAGTAAGAAACCTTTCTCAAAAGAAAGAAACCTTTTGGAACTGTA | 359 |
| Db | 409 | ACTGATGTTTTGACAGTCAGTAAGAAACCTTTCTCAAAAGAAAGAAACCTTTTGGAACTGTA | 468 |
| QY | 360 | CAGCCCTTTTCTTCATTCCTCTTTTGGCTTTCGTGTCGAATGCTTTGGTTCGATGTA | 419 |
| Db | 469 | CAGCCCTTTTCTTCATTCCTCTTTTGGCTTTCGTGTCGAATGCTTTGGTTCGATGTA | 528 |
| QY | 420 | TTATGAAACGTTGATGCGAAGCTTGAGGTTTTTATTTATAGTGGCTTGAAGCTTGG | 479 |
| Db | 529 | TTATGAAACGTTGATGCGAAGCTTGAGGTTTTTATTTATAGTGGCTTGAAGCTTGG | 588 |
| QY | 480 | ATACGTGTTGTTACAGAGATACCCTTATTAAGTTTGGCGAGCTGATGTCCTTA-TTTTT | 538 |
| Db | 589 | ATACGTGTTGTTACAGAGATACCCTTATTAAGTTTGGCGAGCTGATGTCCTTATTTTTT | 648 |
| QY | 539 | TCCCTTTGAAGTACTGAGCCCTCTCTGGTTTTTTTCCCTTGAACAGCTGAGGCTTACAT | 598 |
| Db | 649 | TTCCTTTGAAGTACTGAGCCCTCTCTGGTTTTTTTCCCTTGAACAGCTGAGGCTTACAT | 708 |
| QY | 599 | TTTTCTATGGGATTTTTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT | 658 |
| Db | 709 | TTTTCTATGGGATTTTTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT | 768 |
| QY | 659 | TTTCTAGTAAACATGTTGATTAACCTGGGATTTACATGTTGATTAATACCTGTCATCGT | 718 |
| Db | 769 | TTTCTAGTAAACATGTTGATTAACCTGGGATTTACATGTTGATTAATACCTGTCATCGT | 828 |
| QY | 719 | TTTCTAGTAAACATTAATGCGATTTATAGAAATACGTAATTCCTGATTTCC-TTTTTTTTT | 777 |
| Db | 829 | TTTCTAGTAAACATTAATGCGATTTATAGAAATACGTAATTCCTGATTTCCCTTTTTTTT | 888 |
| QY | 778 | ATCTCTATGCTCTGTGTGTACAGGTCAACAGACCTTCACTCCTATTTTATTTATAGAT | 837 |
| Db | 889 | ATCTCTATGCTCTGTGTGTACAGGTCAACAGACCTTCACTCCTATTTTATTTATAGAT | 948 |
| QY | 838 | TTTATATGCAAGTCGTGCGTGGTTCCTTGGTGTGTAGAGATACAGCCCTTAATTTCTCTGA | 897 |
| Db | 949 | TTTATATGCAAGTCGTGCGTGGTTCCTTGGTGTGTAGAGATACAGCCCTTAATTTCTCTGA | 1008 |
| QY | 898 | GGCATGCTCAGTAAGGCGGGTGTCTCAATGCGGTTCCAAATGATAAAGCGGACGTTTGCT | 957 |
| Db | 1009 | GGCATGCTCAGTAAGGCGGGTGTCTCAATGCGGTTCCAAATGATAAAGCGGACGTTT-CT | 1067 |
| QY | 958 | GCTGCTTCCCGAGATCCAGGACACTAATGCTGTGTC-ACTGAGTATTAATGCTTC | 1018 |
| Db | 1068 | GCTGCTTCCCGAGATCCAGGACACTAATGCTGTGTCAGCTGAGGTAAATGCTTC | 1128 |
| QY | 1017 | AGATCCCAAGAGAGTCAAGATCCAGTGCATATCTTAAAGAAAGATGAATCTTTCTGA | 1076 |
| Db | 1127 | AGATCCCA-AGAAAGTGAATCCAGTGCATATCTTAAAGAAAGATGAATCTTTCTGA | 1185 |
| QY | 1077 | AATATTTTGGCATAGGAAGCAACCTCATGATTTTGTGGAGCTTAATATTTTGGTA | 1138 |
| Db | 1186 | AATATTTTGGCATAGGAAGCAACCTCATGATTTTGTGGAGCTTAATATTTTGGTA | 1245 |
| QY | 1137 | ACGAGTGCATAGGTTTTTAAACACAGCTTCAGCATGCTTAACAGATCACAGGCTTTATGCA | 1198 |
| Db | 1246 | ACGAGTGCATAGGTTTTTAAACACAGCTTCAGCATGCTTAACAGATCACAGGCTTTATGCA | 1305 |
| QY | 1197 | GAACTGATGCTTGATGCTGTTGCAAGCTGTTTACGCACTGCTTGCAGATGACGATTC | 1256 |


```

Db 1306 GAAGTGA-----TGCCTGTGACAGCTGTTACGGCACTGCCCTGCAGTGAAGCATTCG 1357
QY 1257 AGATAGGGGTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1316
Db 1358 AGATAGGGGTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1416
QY 1317 AACACATCTCACCCTGCTGGGTACTTTCAAACCATCTTACAGTACAGTACAGTACAGTAC 1376
Db 1417 AACATCTCACCCTGCTGGGTACTTTCAAACCATCTTACAGTACAGTACAGTACAGTAC 1476
QY 1377 TGAACAGAGAAAGTTCCTCAGTTGATTCATGATGATGATGATGATGATGATGATGATG 1436
Db 1477 TGAACAGAGAAAGTTCCTCAGTTGATTCATGATGATGATGATGATGATGATGATGATG 1536
QY 1437 CAAAGTATATAAGATCTCTATTTGTAATATGACACTGTTAGTCTGTAATCTTT 1496
Db 1537 CAAAGTATATAAGATCTCTATTTGTAATATGACACTGTTAGTCTGTAATCTTT 1596
QY 1497 CTATAGCACCACTTATTTGACAGAGTGTAGGCTGTGGTGGCTGTGGTGGCTGTGGTGG 1556
Db 1597 CTATAGCACCACTTATTTGACAGAGTGTAGGCTGTGGTGGCTGTGGTGGCTGTGGTGG 1656
QY 1557 TCTTTTAA 1564
Db 1657 TCTTTTAA 1664

```

```

RESULT 5
LOCUS Cve277960 4672 bp DNA circular SYN 12-APR-2002
DEFINITION Cloning vector pMAR, luciferase reporter vector containing MAR
insulator sequence.
ACCESSION AJ277960
VERSION AJ277960.1 GI:7981032
KEYWORDS estrogen receptor responsive element; luciferase; MAR insulator;
thymidine kinase promoter.
SOURCE
ORGANISM
Cloning vector pMAR
artificial sequences; vectors.
REFERENCE
1 Clana, P., Di Luccio, G., Belcredito, S., Pollio, G., Vegeto, E.,
Tatangelo, L., Tiverson, C. and Maggi, A.
Engineering of a mouse for the in vivo profiling of estrogen
receptor activity
Mol. Endocrinol. 15 (7), 1104-1113 (2001)
JOURNAL
MEDLINE
21329578
PUBMED
11435611
REFERENCE
2 (bases 1 to 4672)
MAGGI, A.
Direct Submission
Submitted (09-MAY-2000) Maggi A., Institute of Pharmacological
Sciences, University of Milan, Via Balzaretti 9, Milan I-20133,
ITALY

```

```

FEATURES
source
1..4672
/organism="Cloning vector pMAR"
/db_xref="taxon:125631"
/focus="artificial sequence"
/note="artificial sequence"
1..39
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="estrogen receptor responsive elements ERE"
40..79
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic polylinker"
80..247
/organism="Herpes simplex virus unknown type"
/db_xref="taxon:126283"
248..297
/organism="synthetic construct"
source

```

```

/db_xref="taxon:32630"
/note="synthetic polylinker derived from pGL2 basic vector
(Promega)"
298..1950
/organism="Photinus pyralis"
/db_xref="taxon:7054"
2969..3004
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="artificial sequence derived from pGL2basic vector
(Promega)"
3005..4672
/organism="Gallus gallus"
/db_xref="taxon:9031"
/note="MAR insulator sequence from lysozyme gene"
1..39
/bound_moiety="Estrogen receptors"
/vidence="experimental"
80..247
/note="thymidine kinase (TK) promoter"
/vidence="experimental"
298..1950
/organism="Luciferase"
298..1950
/EC_number="1.13.12.7"
/codon_start=1
/product="firefly luciferase"
/protein_id="CAB91857.1"
/db_xref="GI:7981033"
/translation="MEDAKNIKRPAPFPYLEDGTAGEOLHKAMKRYALVPGTIAFTD
AHEVNIITAEYEMSVRLAEAMKRYGNRIYVCSKNSLOEPMPVGALEIVAV
APANDVYNERELNSMNTISOPVYVPSKSKLOKILNVOKRIPILOKILIMDSKTDVOC
FOSMTFVTSHPPEFNEYDEVPESDRKXTALINSSGSGTLPBGVALPHRTACVR
FSHARDPIFGNOIIPDTALISVPEFHGGMFTLGYLLCGFRVVMYMPFEELFHS
LDDYKIOSALVPLETSEFPAKSTLIDKXDSNHEIASGSAPLSKRGVAVKREFLP
GJROGGLTETTSAILITPEDDKPGAVKVPFPAKYVDLDGTGLGNORGELCV
RGMIMSGYVNNDEATNALIDKDGMSGSDIAYWDEDEHFTVDRKSLIKYKGYOVA
PALESIILOHNPIDPAGVAGLPDDAGELPAVAVYVLEHKGKTEIYDYVASOYTT
AKLRGVFVDEVPKGLTKIDARKIREILIKAKGSKL"
BASE COUNT 1262 a 916 c 1056 g 1438 t
ORIGIN
Query Match 11.5%; Score 1374.4; DB 12; Length 4672;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1523; Conservative 0; Mismatches 16; Indels 29; Gaps 11;
QY 1 TGCCGCTCTTGTGATTCACCTGTTGTATTTCATCTCTCTGCGGATGAAGGATA 60
|||||
Db 3126 TGCCTCTTGTGATTCACCTGTTGTATTTCATCTCTCTGCGGATGAAGGATA-- 3183
QY 61 TAAACAGTCTGTATACAGTCTGTGAGGAATCTGTGATTTCTTGTGATCAGTGT 120
|||||
Db 3184 -----TATAACAGTCTGTGAGGAATCTGTGATTTCTTGTGATCAGTGT 3233
QY 121 ATAAAGTATGTGATTTGTGATTAAGGCTGTGCTCTTGTCTGGGAGACAAGCCCA 180
|||||
Db 3234 ATAAAGTATGTGATTTGTGATTAAGGCTGTGCTCTTGTCTGGGAGACAAGCCCA 3293
QY 181 CAGCAGTGTGTTGGGTGGGTGGGACGACGACGACGACGACGACGACGACGACGACG 240
|||||
Db 3294 CAGCAGTGTGTTGGGTGGGTGGGACGACGACGACGACGACGACGACGACGACGACG 3352
QY 241 TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT 299
|||||
Db 3353 TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT-TTTTGT 3412
QY 300 ACTGATGTTTGTGACAGTGAGAACATTTCTCAAAAGAAAGACCTTTGGAACGTGA 359
|||||
Db 3413 ACTGATGTTTGTGACAGTGAGAACATTTCTCAAAAGAAAGACCTTTGGAACGTGA 3472
QY 360 CAGCCCTTTCTTCAATCCCTTTTGTCTGTGTGCAATGCTTGTGTTGATTCGA 419

```

```

Db 3473 CAGCCCTTTCTTCAATCCCTTTGCTCTGTCGCAATCCCTTGGTCTGATTCGA 3532
Oy 420 TTATGGAAAACGTTGATCGGAACCTGTTTATTTATTTAGTGGCTTGAAGCTTGG 479
Db 3533 TTATGGAAAACGTTGATCGGAACCTGTTTATTTATTTAGTGGCTTGAAGCTTGG 3592
Oy 480 ATAGCTGTTGTTACACGAGATACCTTATTAAGTTAGCCAGCTGATCCTTTA-TTTTT 538
Db 3593 ATAGCTGTTGTTACATGATACCTTATTAAGTTAGCCAGCTGATCCTTATTTATTT 3652
Oy 539 TCCCTTTGAAGTAGAGGCTCTCGTGTGTTTTCCTTTGAACGGTAGGCTTAGAT 598
Db 3653 TTCTTTGAAGTAGAGGCTCTCGTGTGTTTTCCTTTGAACGGTAGGCTTAGAT 3712
Oy 599 TTTTCTAATGGGATTTTTCCTGATGATCTAGTTGATACCCAAATGCTTGAATGTT 658
Db 3713 TTTTCTAATGGGATTTTTCCTGATGATCTAGTTGATACCCAAATGCTTGAATGTT 3772
Oy 659 TTCTAGTTAATCATGTGTGATTAACCTTCCGATTTACATGTTGTATATCTGATCTGT 718
Db 3773 TTCTAGTTAATCATGTGTGATTAACCTTCCGATTTACATGTTGTATATCTGATCTGT 3832
Oy 719 TTCTAGTTAATCATGTGTGATTAACCTTCCGATTTACATGTTGTATATCTGATCTGT 777
Db 3833 TTCTAGTTAATCATGTGTGATTAACCTTCCGATTTACATGTTGTATATCTGATCTGT 3892
Oy 778 ATCTGTATGCTCTGTGTATGATGATCAAGCTCAACAGACTTCACTCTATTTTATTTAGAT 837
Db 3893 ATCTGTATGCTCTGTGTATGATGATCAAGCTCAACAGACTTCACTCTATTTTATTTAGAT 3952
Oy 838 TTTATATGAGCTGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
Db 3953 TTTATATGAGCTGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4012
Oy 898 GCGATGCTCAGTAAGCGCGGTGTGATCATGATGATGATGATGATGATGATGATGATGAT 957
Db 4013 GCGATGCTCAGTAAGCGCGGTGTGATCATGATGATGATGATGATGATGATGATGATGAT 4071
Oy 958 GCTGCTTCCGAGATCCAGACACTAACTCTCTCTG-ACGTAGATTAATGCTTGC 1016
Db 4072 GCTGCTTCCGAGATCCAGACACTAACTCTCTCTG-ACGTAGATTAATGCTTGC 4130
Oy 1017 ACATCCGAGGAGGAGGAGATCCAGATGATGATGATGATGATGATGATGATGATGATGAT 1076
Db 4131 ACATCCGAGGAGGAGGAGATCCAGATGATGATGATGATGATGATGATGATGATGATGAT 4189
Oy 1077 AATATTTTGGCATAGGAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1136
Db 4190 AATATTTTGGCATAGGAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4249
Oy 1137 ACGAGTGCATAGGTTTTAAACACAGTTGCAGATGCTAACAGAGTCACAGCTTTATGCA 1196
Db 4250 ACGAGTGCATAGGTTTTAAACACAGTTGCAGATGCTAACAGAGTCACAGCTTTATGCA 4309
Oy 1197 GAAAGTATGCTGAGATGCTGTGTGCAAGCTGTTTACGCACTGCTTGCAGTAGCATTCG 1256
Db 4310 GAAAGTATGCTGAGATGCTGTGTGCAAGCTGTTTACGCACTGCTTGCAGTAGCATTCG 4361
Oy 1257 AGATGGGGTGGGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1316
Db 4362 AGATGGGGTGGGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4420
Oy 1317 AACACATCTCACTGCTGGTACTTTTCAAAACATCTTAGCAGTAGTAGTAGTAGTAGTAGTA 1376
Db 4421 AACACATCTCACTGCTGGTACTTTTCAAAACATCTTAGCAGTAGTAGTAGTAGTAGTAGTA 4480
Oy 1377 TGAACAAGAGAGTTCCTCAGTGTGATATCTCATGAGGATGCTTTTTTCCATGTTGGG 1436
Db 4481 TGAACAAGAGAGTTCCTCAGTGTGATATCTCATGAGGATGCTTTTTTCCATGTTGGG 4540
Oy 1487 CAAAGTATGATTAAGCATCTCATATTGTAATATGCACTGTGTAGTCCATACTCTT 1496

```

```

Db 4541 CAAAGTATGATTAAGCATCTCATATTGTAATATGCACTGTGTAGTCCGTAATCTTT 4600
Oy 1497 CTATACACACACTTATTTGACAGAGGTAGGCTGTGTGGCTGTGTGCTGTGCTCA 1556
Db 4601 CTATACACACACTTATTTGACAGAGGTAGGCTGTGTGGCTGTGTGCTGTGCTCA 4660
Oy 1557 TCTTTTAA 1564
Db 4661 TCTTTTAA 4668

RESULT 6
CLCURVAS
LOCUS CLCURVAS 660 bp DNA linear VRT 22-OCT-1990
DEFINITION Chicken lysozyme gene intrinsically curved segment of DNA.
ACCESSION X52989
VERSION X52989.1 GI:62716
KEYWORDS curved DNA; matrix attachment region.
SOURCE
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 660)
REFERENCE
AUTHORS Straetling, W.H.
TITLE Direct Submision
JOURNAL Submitted (10-MAY-1990) Straetling W.H., Physiologisch-Chemisches
Institut, Universitaets-Krankenhaus Eppendorf, Martinistr. 52,
D-2000 Hamburg, FRG
2 (bases 1 to 660)
REFERENCE
AUTHORS von Kries, J.P., Phil-Van, L., Diekmann, S. and Straetling, W.H.
TITLE A non-curved chicken lysozyme 5' matrix attachment site is 3'
followed by a strongly curved DNA sequence
JOURNAL Nucleic Acids Res. 18 (13), 3881-3885 (1990)
MEDLINE 90326509
PUBMED 2374712
COMMENT
Data kindly reviewed (11-AUG-1990) by Straetling W. This sequence
represents an intrinsically curved segment of the chicken lysozyme
5' matrix attachment region (MAR).
FEATURES
source
1..660
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="lambda-Lys4/2."
BASE COUNT 179 a 123 c 126 g 232 t
ORIGIN
Query Match 5.4%; Score 644.2; DB 5; Length 660;
Best Local Similarity 99.4%; Pred. No. 1.8e-141;
Matches 657; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy 2011 GCGCTGCTGACTTTTAAACATTAAGCTGATTTGCTTCTGATGATGATGATGATGATGATGAT 2070
Db 1 GCGCTGCTGACTTTTAAACATTAAGCTGATTTGCTTCTGATGATGATGATGATGATGATGAT 60
Oy 2071 TTCCCAATTTGCAACAAGATGCTGCGTAACATTTCAAGAAATGCTTGAATATACAC 2130
Db 61 TTCCCAATTTGCAACAAGATGCTGCGTAACATTTCAAGAAATGCTTGAATATACAC 120
Oy 2131 ATGGAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2190
Db 121 ATGGAGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Oy 2191 CTCTCAGAAATGCCCAACATCCCAAGAGATTTATATGATATATAGTATAGCAGTTTCTGAT 2250
Db 181 CTCTCAGAAATGCCCAACATCCCAAGAGATTTATATGATATATAGTATAGCAGTTTCTGAT 239
Oy 2251 TCCACAGGCCCAAGAGATCTGCTGAATGTTGTGTGCGGAGACCTGATATTTCTCAACAA 2310
Db 240 TCCACAGGCCCAAGAGATCTGCTGAATGTTGTGTGCGGAGACCTGATATTTCTCAACAA 239
Oy 2311 GGTAGATGATATCTAGCACTGCGGATTTTAAATACATTTTTCACAGACAGTACTAGTT 2370

```

Db 300 GGTAAGATGTATTCCTAGACACGCGGATTTAATACATTTTCAGCAGAACTAGT 359
QY 2371 AATCTCTACCTTAAAGGATCGTTTCATCTTTTATAGATGTATTAATCTGCATA 2430
Db 360 AATCTCTACCTTAAAGGATCGTTTCATCTTTTATAGATGTATTAATCTGCATA 419
QY 2431 ACTTTTACCTTCANAGGTCCTTTTTCAGCCTTTAGAGACGTGTAAACAATTGCT 2490
Db 420 ACTTTTACCTTCANAGGTCCTTTTTCAGCCTTTAGAGACGTGTAAACAATTGCT 479
QY 2491 GTCCAACTTTGTGTGTCTTAACTGCAATAGTATGTTACCTGTATGAAGAAATA 2550
Db 480 GTCCAACTTTGTGTGTCTTAACTGCAATAGTATGTTACCTGTATGAAGAAATA 539
QY 2551 AGACCACTTTTATATTAATAAATACCTTTGTCTGCTCATTTTGTGCTGATATCC 2610
Db 540 AGACCACTTTTATATTAATAAATACCTTTGTCTGCTCATTTTGTGCTGATATCC 599
QY 2611 TTGCAGTGGCCATTTATGTCAGTTCTGTGCAATATTCAGACATCAAACTTAACGTAGCT 2670
Db 600 TTGCAGTGGCCATTTATGTCAGTTCTGTGCAATATTCAGACATCAAACTTAACGTAGCT 659
QY 2671 C 2671
Db 660 C 660

RESULT 7

CHKLSZF

LOCUS 720 bp DNA linear VRT 13-MAY-1999
DEFINITION Gallus gallus egg white lysozyme gene, partial cds.
ACCESSION J00886 V00429
VERSION J00886.1 GI:212277

KEYWORDS

SOURCE

Gallus gallus.
Gallus gallusORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.REFERENCE
1 (bases 1 to 720)
Grez,M., Land,H., Glesecke,K., Schutz,G., Jung,A. and Sippel,A.E.TITLE
Multiple mRNAs are generated from the chicken lysozyme gene
JOURNAL
Cell 25 (3), 743-752 (1981)
MEDLINE
82025586
PUBMED
7285117

REFERENCE

AUTHORS

2 (bases 220 to 440)
von der Ahe,D., Janich,S., Scheiderreit,C., Renkawitz,R., Schultze,G.TITLE
and Beato,M.
Glucocorticoid and progesterone receptors bind to the same sites in
two hormonally regulated promotersJOURNAL
Nature 313 (6004), 706-709 (1985)
MEDLINE
85137874
PUBMED
2983219

REFERENCE

AUTHORS

3 (sites)
Wolff,S., Schrader,M. and Wittig,B.TITLE
Lack of correlation between DNA methylation and transcriptional
inactivation: the chicken lysozyme gene
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 88 (1), 271-275 (1991)
MEDLINE
91095443
PUBMED
1986375COMMENT
On May 13, 1999 this sequence version replaced gi:63584.
FEATURES
source

FEATURES

1..720
Location/Qualifiers/organism="Gallus gallus"
/db_xref="taxon:9031"

mRNA

440..>604
/product="egg white lysozyme"

exon

440..604
/number=1
469..>604
/codon_start=1

CDS

/product="egg white lysozyme"
/protein_id="AAD10202.1"

/db_xref="GI:4204093"
/translation="MSRLIIIVLCFLPLAALGKVRGRCLEAAMKRHGLDNRGSLG
N"
BASE COUNT 202 a 126 c 186 g 206 t
ORIGIN

Query Match 4.3%; Score 515.6; DB 5; Length 720;
Best local similarity 99.2%; Pred. No.5.9e-111;
Matches 518; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 11424 TGGTCAGATGCTGTTTATATAGCAGATATGACCAAGTTGCACCTATGATGACAAACA 11483
Db 1 TGGTCAGATGCTGTTTATATAGCAGATATGACCAAGTTGCACCTATGATGACAAACA 60

QY 11484 ATGTGTGTCATCCTTAGCAGCTGAGAAAGAGCCAAATTTGCATTTGTATAGAAATGCT 11543
Db 61 ATGTGTGTCATCCTTAGCAGCTGAGAAAGAGCCAAATTTGCATTTGTATAGAAATGCT 120

QY 11544 TTAGTAATTCCTGCCAATTAACCTTGTATATACATGAGCTGTTTATGCTGTAGT 11603
Db 121 TTAGTAATTCCTGCCAATTAACCTTGTATATACATGAGCTGTTTATGCTGTAGT 180

QY 11604 AGTGTAACACTGATGATGAACAATGGCTATGCACTAAATCAAGACTGTATATGCAA 11663
Db 181 AGTGTAACACTGATGATGAACAATGGCTATGCACTAAATCAAGACTGTATATGCAA 240

QY 11664 CAGACTATTAATAATTCCTGTGGCTTAGCCAGTGTGCTACTTCCCATTTATAGAAA 11723
Db 241 CAGACTATTAATAATTCCTGTGGCTTAGCCAGTGTGCTACTTCCCATTTATAGAAA 300

QY 11724 TTGGCAAGTTTATAGCAATGTTTGAAGTGTGGAAATTTCTGTATCTCAAGAGGCG 11783
Db 301 TTGGCAAGTTTATAGCAATGTTTGAAGTGTGGAAATTTCTGTATCTCAAGAGGCG 360

QY 11784 TTTTGGACAACCTGTAGAACAGAGCAATCAAAAGGGGTGGAGAACTTAAAGAGAGG 11843
Db 361 TTTTGGACAACCTGTAGAACAGAGCAATCAAAAGGGGTGGAGAACTTAAAGAGAGG 420

QY 11844 CAGGTGCAAGAGAGCTTGCAGTCCCGCTGTGTAGACACTGGCAACATGAGGCTTTTG 11903
Db 421 CAGGTGCAAGAGAGCTTGCAGTCCCGCTGTGTGTAGACACTGGCAACATGAGGCTTTTG 480

QY 11904 CTATCTTGGTGTCTTGTGCTTCTGCCCCGTGCTGCTTATAGG 11945
Db 481 CTATCTTGGTGTCTTGTGCTTCTGCCCCGTGCTGCTGCTGCTG 522

RESULT 8

GGDNALGP

LOCUS 449 bp DNA linear VRT 11-FEB-1995
DEFINITION G.gallus lysozyme gene promoter.
ACCESSION X84223
VERSION X84223.1 GI:666995

KEYWORDS

ARBP binding site; promoter.
SOURCE
chicken.

ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.REFERENCE
1 (bases 1 to 449)
Buhmester,H., von Kries,J.P. and Straetling,W.H.TITLE
Nuclear matrix protein ARBP recognizes a novel DNA sequence motif
with high affinity
JOURNAL
Biochemistry
2 (bases 1 to 449)
Straetling,W.H.
Submitted (27-JAN-1995) W.H. Straetling, Institut fuer
Physiologische Chemie, Universitaets-Krankenhaus Eppendorf,
Martinistrasse 52, 20246 Hamburg, FRGJOURNAL
Biochemistry
2 (bases 1 to 449)
Straetling,W.H.
Submitted (27-JAN-1995) W.H. Straetling, Institut fuer
Physiologische Chemie, Universitaets-Krankenhaus Eppendorf,
Martinistrasse 52, 20246 Hamburg, FRG

AUTHORS

2 (bases 1 to 449)
Straetling,W.H.
Submitted (27-JAN-1995) W.H. Straetling, Institut fuer
Physiologische Chemie, Universitaets-Krankenhaus Eppendorf,
Martinistrasse 52, 20246 Hamburg, FRG

JOURNAL

2 (bases 1 to 449)
Straetling,W.H.
Submitted (27-JAN-1995) W.H. Straetling, Institut fuer
Physiologische Chemie, Universitaets-Krankenhaus Eppendorf,
Martinistrasse 52, 20246 Hamburg, FRG

AUTHORS

2 (bases 1 to 449)
Straetling,W.H.
Submitted (27-JAN-1995) W.H. Straetling, Institut fuer
Physiologische Chemie, Universitaets-Krankenhaus Eppendorf,
Martinistrasse 52, 20246 Hamburg, FRG

JOURNAL

2 (bases 1 to 449)
Straetling,W.H.
Submitted (27-JAN-1995) W.H. Straetling, Institut fuer
Physiologische Chemie, Universitaets-Krankenhaus Eppendorf,
Martinistrasse 52, 20246 Hamburg, FRG

FEATURES

source
1..449
Location/Qualifiers
/organism="Gallus gallus"

/db_xref="taxon:9031"
1..449
/note="portion of the MAR attachment region of lysozyme gene"
protein_bind 111..171
/bound_molety="nuclear binding protein ARBP"
protein_bind 311..371
/bound_molety="nuclear binding protein ARBP"
BASE COUNT 157 a 71 c 76 g 145 t
ORIGIN

Query Match 3.6%; Score 432.8; DB 5; Length 449;
Best Local Similarity 98.9%; Pred. No. 2.5e-91;
Matches 447; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 1564 AAGCTTTTGGAAATACACTGATTTGATTAAGTCTTTGAAGATAGTAACACTACTT 1623
|||||
Db 1 AAGCTTTTGGAAATACACGACTTGAAGTCTTTGAAGATAGTAACACTACTT 60
|||||

QY 1624 ACCTTTGATCCCAATGAATGAGCATTTCAGTTGTAAGAATTCGGCTATTGATACC 1683
|||||
Db 61 ACCTTTGATCCCAATGAATGAGCATTTCAGTTGTAAGAATTCGGCTATTGATACC 120
|||||

QY 1684 ATGTAATGTAATTTACACCCCACTGCTGACACTTTGGAATATTTCAAGTAATAGACT 1743
|||||
Db 121 ATGTAATGTAATTTACACCCCACTGCTGACACTTTGGAATATTTCAAGTAATAGACT 180
|||||

QY 1744 TTGGCTCACCCTCTTGTGCTACTGATTTTGTAAATAGAAATATTTTAACTGTCATAT 1803
|||||
Db 181 TTGGCTCACCCTCTTGTGCTACTGATTTTGTAAATAGAAATATTTTAACTGTCATAT 240
|||||

QY 1804 GATTATTACATTATGAAGAAGACATTCGTCGATCTTCAATGTAGAAGAAATGAGAGATG 1863
|||||
Db 241 GATTATTACATTATGAAGAAGACATTCGTCGATCTTCAATGTAGAAGAAATGAGAGATG 300
|||||

QY 1864 CGTGTGCTTTTATTAATCAAGTGTGCAATTTAGTGAGGTGCTTAAAGAAAAA 1923
|||||
Db 301 CGTGTGCTTTTATTAATCAAGTGTGCAATTTAGTGAGGTGCTTAAAGAAAAA 357
|||||

QY 1924 AAAAAAGTAATATATAAAGAGACAGGTGTTTACAAAGTAATATTCCTATTGTTGTA 1983
|||||
Db 358 AAAAAAGTAATATATAAAGAGACAGGTGTTTACAAAGTAATATTCCTATTGTTGTA 417
|||||

QY 1984 ACAGTTACATTTTATGAAGATTACACAGCGT 2015
|||||
Db 418 ACAGTTACATTTTATGAAGATTACACAGCGT 449
|||||

RESULT 9
GLYSHRE 441 bp DNA linear VRT 06-JUL-1989
LOCUS
DEFINITION Chicken lysozyme gene promoter region with hormone responsive
element (HRE).
ACCESSION X12509
VERSION X12509.1 GI:63583
KEYWORDS hormone receptor; lysozyme; steroid hormone receptor.
SOURCE
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 441)
Hecht,A., Berkenstam,A., Stromstedt,P.E., Gustafsson,J.A. and
Slipfel,A.E.
TITLE
A progesterone responsive element maps to the far upstream steroid
dependent DNase hypersensitive site of chicken lysozyme chromatin
EMBO J. 7 (7), 2063-2073 (1988)
JOURNLT
MEDLINE 88328995
PUBMED 3416833
COMMENT
author numbering according transcript start site -2250 to -1815;
see V00429 for lysozyme gene exon 1.
FEATURES
Data kindly reviewed (17-Apr-1989) by Berkenstam A.
Location/Qualifiers

source 1..441
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="lambda lys30 and lambda lys31."
misc_feature 122..138
/note="glucocorticoid receptor site"
misc_feature 122..138
/note="progesterone receptor site"
misc_feature 151..169
/note="progesterone receptor site"
misc_feature 153..169
/note="glucocorticoid receptor site"
misc_feature 188..220
/note="glucocorticoid receptor site"
misc_feature 189..220
/note="progesterone receptor site"
misc_feature 254..295
/note="progesterone receptor site"
misc_feature 265..293
/note="glucocorticoid receptor site"
misc_feature 298..374
/note="glucocorticoid receptor site"
misc_feature 298..363
/note="progesterone receptor site"
BASE COUNT 122 a 107 c 108 g 104 t
ORIGIN

Query Match 3.5%; Score 421; DB 5; Length 441;
Best Local Similarity 98.6%; Pred. No. 1.6e-88;
Matches 435; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 9621 ACTACTTTTGCAGTAATCATAGGGTTAGTATGTTAATCTGA-GGAAAAA 9679
|||||
Db 1 ACTACTTTTGCAGTAATCATAGGGTTAGTATGTTAATCTGAAGGAAAAA 60
|||||

QY 9680 GCCAACCCGACAGACATCCAGCTCAGGTGGAATCAAGATCAAGTCAGTCAGTGGGTC 9739
|||||
Db 61 GCCAACCCGACAGACATCCAGCTCAGGTGGAATCAAGATCAAGTCAGTCAGTGGGTC 120
|||||

QY 9740 CCAGAGAACACAGGAGACTCTTCTTTAGAGCTTTATGTACAGGGCCCAAGATACTGA 9799
|||||
Db 121 CCAGAGAACACAGGAGACTCTTCTTTAGAGCTTTATGTACAGGGCCCAAGATACTGA 180
|||||

QY 9800 TCTTGTAGTGAAGAGACTTTCATTTGCGCACAGTTGACGTGAGGCAATCTGGAATTTTC 9859
|||||
Db 181 TCTTGTAGTGAAGAGACTTTCATTTGCGCACAGTTGACGTGAGGCAATCTGGAATTTTC 240
|||||

QY 9860 TCTCGCTGCACAGTTCCAGTATCCAGTATGTTGACAGTTTGGGACACTTTTGGGTCAGG 9919
|||||
Db 241 TCTCGCTGCACAGTTCCAGTATCCAGTATGTTGACAGTTTGGGACACTTTTGGGTCAGG 300
|||||

QY 9920 CCGTATCCCAAGAGCAGAGAGTTCCAGCTATGTCAGGAGAGTCCCTGACCCCACTC 9979
|||||
Db 301 CCGTATCCCAAGAGCAGAGAGTTCCAGCTATGTCAGGAGAGTCCCTGACCCCACTC 360
|||||

QY 9980 ACTGCATCAACAAGGCGAAGACCAAGAGATGGCTTTTGTGAATTTGCAGTGTGGCC 10039
|||||
Db 361 ACTGCATCAACAAGGCGAAGACCAAGAGATGGCTTTTGTGAATTTGCAGTGTGGCC 420
|||||

QY 10040 CAGAGGGGCTGCACCACTACT 10060
|||||
Db 421 CAGAGGGGCTGCACCACTACT 441
|||||

RESULT 10
CHKPOLLI/c 2232 bp DNA linear VRT 15-OCT-1993
LOCUS
DEFINITION Gallus gallus reverse transcriptase gene, 3' end.
ACCESSION L22152
VERSION L22152.1 GI:347226
KEYWORDS POL-1 like gene.
SOURCE
ORGANISM Gallus gallus
Gallus gallus (strain White Leghorn, sub-species domesticus) DNA.

NKGLVTVGYVYRPDPDQDEAFELVLMLESCQALVLEFNPHRIICMDFDHLYKE
 RSRRLVBCIDDDNFLVQVDRPTRECVLDDLTITAAEILKGVNNGSLGCSGDHALYKE
 MISRDVGLANGVATLTFNGRENEFKLGSJLAMIPTWNVIKDQVDEESMLERKDALILKRA
 QEVSILPLNTRVGRDRDPAMVIGKDLITLCTLRKICAKYALKMGQITREWEYRAVATCR
 DTRVAVQVNIAPRUKNKKRVEPVITGVDPQKRVNGLDVIHFKCELAQENKATK

| Db | 535 | ATGATGAGCCAGCATACATTGGCTTTCCGAGCTGCA | 496 |
|-----------------------|---|--------------------------------------|-----|
| RESULT 13 | | | |
| CHRPOLIC/c | | | |
| LOCUS | CHRPOLIC | 1651 bp | DNA |
| DEFINITION | Gallus gallus POL-like gene, 3' end. | | |
| ACCESSION | L22148 | | |
| VERSION | L22148.1 | GI:347218 | |
| KEYWORDS | POL-like gene. | | |
| SOURCE | Gallus gallus (strain White Leghorn, sub-species domesticus) DNA. | | |
| ORGANISM | Gallus gallus | | |
| REFERENCE | 1 (bases 1 to 1651) | | |
| AUTHORS | Burch,J.B., Davis,D.L. and Haas,N.B. | | |
| TITLE | Chicken repeat 1 elements contain a pol-like open reading frame and belong to the non-long terminal repeat class of retrotransposons | | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 90 (17), 8199-8203 (1993) | | |
| MEDLINE | 93376771 | | |
| PUBMED | 8396264 | | |
| FEATURES | | | |
| source | location/qualifiers | | |
| | 1..1651 | | |
| | /organism="Gallus gallus" | | |
| | /strain="White Leghorn" | | |
| | /sub_species="domesticus" | | |
| | <1..381 | | |
| | /note="POL-like" | | |
| | /codon_start=1 | | |
| | /label=ORF | | |
| | /protein_id="AA049023.1" | | |
| | /db_xref="GI:347218" | | |
| | /translation="LCTSPGPGGOGKSPPTVAVAEVDRDLRLDEYKSGPDGVHP RVLEKEVVAEPLSTFEKSNLSGSEVPDDMRKQVYPIPKKSKEDPGNRPVLSLS VPGKIMQIILDDMLDHMRNRY" | | |
| CDS | | | |
| | /note="POL-like" | | |
| | /codon_start=1 | | |
| | /label=ORF | | |
| | /protein_id="AA049023.1" | | |
| | /db_xref="GI:347218" | | |
| | /translation="LCTSPGPGGOGKSPPTVAVAEVDRDLRLDEYKSGPDGVHP RVLEKEVVAEPLSTFEKSNLSGSEVPDDMRKQVYPIPKKSKEDPGNRPVLSLS VPGKIMQIILDDMLDHMRNRY" | | |
| BASE COUNT | 405 a 360 c 531 g 355 t | | |
| ORIGIN | | | |
| Query Match | 3.2%; Score 377.6; DB 5; Length 1651; | | |
| Best Local Similarity | 79.4%; Pred. No. 3.4e-78; | | |
| Matches | 508; Conservative 0; Mismatches 104; Indels 28; Gaps 4; | | |
| OY 10574 | ATGAGTATATGAGTATGAGCTGGCTGGTGGAGAGGACCCCAAGATCATGAGATCCACACC | 10633 | |
| Db 1651 | ACAGATATACAGAAATGACCTGGGTGGATGGACTCAAGGATCATGATCTCCAAACCT | 1592 | |
| OY 10634 | CCCGCACAGGAGGAGGCCACCAACCTCCAGATCTGTACTAGACGAGGAGCGGCT | 10693 | |
| Db 1591 | ACTGC--CGTGACAGGAGGCCACCAACCTCCCAT--TTACTAGACGAGGCTGCCAGGCTC | 1537 | |
| OY 10694 | CCATCAACCTGGCCCATGAGACACCTCCAGAGGATGAGGAGCATCCAAACCTCTCTGGGACG | 10753 | |
| Db 1536 | TCATCCACACCTGGGCTTGAACACCTCCAGGAGTGGGCGATCCAAACCTCTCTGGGACG | 1477 | |
| OY 10754 | CTGTGCGAGACCTACCAACCTCTCTGTGGAAGACTTTCCCTACATCCATATAGC | 10813 | |
| Db 1476 | CCATTCACGACCTACCAACCTCTCTGTGTAAGACTTCCCTATATCCAACTTAAT | 1417 | |
| OY 10814 | CTTCCCTCTGAGGTATGATCCATCTCCCTCTGTGATCATCTGTACTCTTTTAA | 10873 | |
| Db 1416 | CTTCCCTCTTTCACCTTAAACCGTTCGCTGTCTGTATCTACCTTTCAAG | 1357 | |
| OY 10874 | AGTTGATCTCTCTCTTT-----TGAAGGTTGCAATGAGGCT | 10913 | |
| Db 1356 | AGTTGATCTCTCTCTTTATATAGCTCCCTTAGATCTGAAGGCGCATGAGGTCA | 1297 | |
| OY 10914 | CTTTCAGCCTTCTCTCTGAGGATGAACAGACGACGCTCCCTGAGCTGCTTAA | 10973 | |
| Db 1296 | CCCTGACG---CTTCTCTTCTCCAGGCTGAATGAAGCCCACTCCCTTACGCTGCTCG | 1240 | |
| OY 10974 | TAGGAGAGGTGCTCCAGCCTCTGTATCATCTTTGTGGCCCTCTCTGAGACCGGCTCAAG | 11033 | |

| | | | |
|---------------|---|--|--------------------------------|
| Db | 1239 | CAGGGAGAGTGCCTCAGGCCCTTATTCCTTTTGGGCCCTCTGTGGACCTCTTCACG | 1180 |
| QY | 11034 | AGCTCCACATCTTTCCTGCTTACTGCGGGGCCCAAGGCGCTGATATCAGTACTCCAGATGGGGC | 11093 |
| Db | 1179 | AGCTCTCATCTTCTTCTTACTGAGGAGCTCATACCTGACAGACAGTACTCCAGATGGGGC | 1120 |
| QY | 11094 | CTCAAAAGAGGAGTAAGAGGGGACATCACTTCCTCACCCTGCGTGGCCAGGCGCTTT | 11153 |
| Db | 1119 | CTCACAAAGAGAGATGAGGAGGAGACATCACTTCCTGCTCGCTGCGCCACTCTCTC | 1060 |
| QY | 11154 | CTGATGAGAGCCCTGGATACAACTGGCTTTCAGCTGCA | 11193 |
| Db | 1059 | ATGATGAGAGCCAGATACCATTTCTTTCGAGCTGCA | 1020 |
| RESULT 14 | | | |
| CHKCR1A/C | | | |
| LOCUS | | | |
| DEFINITION | CHKCR1A | 971 bp | DNA |
| | Gallus gallus vitellogenin gene, intron 2 with a CRI retroposon insertion. | | linear VRT 03-JAN-1995 |
| ACCESSION | M28069 | | |
| VERSION | M28069.1 | GI:341819 | |
| KEYWORDS | retroposon; vitellogenin. | | |
| SOURCE | Gallus gallus DNA. | | |
| ORGANISM | Gallus gallus | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Gallus; Phasianinae; Gallus. | | |
| REFERENCE | 1 (bases 1 to 971) | | |
| AUTHORS | Silva,R. and Burch,J.B. | | |
| TITLE | Evidence that chicken CRI elements represent a novel family of retroposons | | |
| JOURNAL | Mol. Cell. Biol. 9 (8), 3563-3566 (1989) | | |
| MEDLINE | 90014816 | | |
| PUBMED | 2477689 | | |
| FEATURES | | | |
| source | Location/Qualifiers | | |
| | 1..971 | | |
| | /organism="Gallus gallus" | | |
| | /db_xref="taxon:9031" | | |
| gene | 1..971 | | |
| | /gene="VTGIII" | | |
| | 1..>971 | | |
| | /gene="VTGIII" | | |
| intron | 105..110 | | |
| | /note="5' insertion target sequence" | | |
| repeat_region | /rpt_type=direct | | |
| | complement(111..946) | | |
| | /note="retroposon" | | |
| | /rpt_family="CRI" | | |
| CDS | <111..853 | | |
| | /gene="VTGIII" | | |
| | /note="start; putative" | | |
| | /codon_start=3 | | |
| | /product="unknown protein" | | |
| | /protein_id="AA58720.1" | | |
| | /db_xref="GI:609434" | | |
| | /translation="RDIDRLERARVNLIRENTAKCRYLHLGRCNPRLHRLKLAAYLE | | |
| | SSSAEKDGLVLMDBKLMSQDCTLAQKANGILISIMKVASRDETIIVPLYSALVRPP | | |
| | HLKCVQVWRPOYKKRDVLLERVDRTEEMIKGIEHPYEDRLRELGLFSLKRLRLG | | |
| | DLIAFQSLKGAIKQEGSKLFEVDNSRTFRNGFKLKEGRFLDVRGKFTFRVVRQW | | |
| | NRLPREVDNLISLEVFARLDGALNLY | | |
| | 947..952 | | |
| repeat_region | /note="3' insertion target sequence" | | |
| | /rpt_type=direct | | |
| BASE COUNT | 233 a | 309 g | 221 t |
| ORIGIN | | | |
| | Query Match | 3.1%; | Score 375.2; DB 5; Length 971; |
| | Best Local Similarity | 77.9%; | Pred. No. 1.2e-77; |
| | Matches 514; Conservative | 0; Mismatches 118; Indels 28; Gaps 4; | |
| Y | 10554 | AAACCTACTTTATCAATCATAGATCATAGATGAGGCTGGGTGGAGAGGACCCCAAG | 10613 |

```

Db 966 AAAAATAATGACAGAAATGACAGAAATGACAGAAATGAGCTGGGTGGAGGAGGACCTCGAG 907
Oy 10614 GATCATGAAAGATCCAAACACCCCGCCACAGAGGGGCCAACCAACCTCCAGATCTGCTACT 10673
Db 906 GATTCAGATCTCTCCAAACCCCTGCG--TGGGCAAGGGCCACCAACATCCCAT--TTACT 852
Oy 10674 AGACGAGGAGCCGAGGAGGCTCCATCCAACTGGCCATGAAACCTCCAGAGGATGAGCAT 10733
Db 851 AGACAGAGTGGCCAGGGGCCCATCCAACTGGCTTGAAACCTCCAGAGGATGAGCAT 792
Oy 10734 CCACAACTCTCTGGGCAAGCTGTGCGACAGACCTCACCACCTCTCTGTGAAGAACTTTT 10793
Db 791 CCACAACTCTCTGGGCAAGCTGTGCGACAGACCTCACCACCTCTCTGTGAAGAACTTTT 732
Oy 10794 CCCTGACATCCAAATGAGCTCCCTGATGAGGTTAGATGATGATGATGATGATGATGAT 10853
Db 731 CCCTGACATCCAAATGAGCTCCCTGATGAGGTTAGATGATGATGATGATGATGATGAT 672
Oy 10854 CACTGTCTACTCTGTGTAAGAAAGTGTATCTCTCTCTT-----T 10893
Db 671 TGTATCTACCTCTTCAAGAGTTTACTCCCTCTCTGTTATAGGCTCCCTTCAGGAGT 612
Oy 10894 GGAAGGTTCAGATGAGTCTCTCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10933
Db 611 GAAAGGCTGCAATGGGATCAGCGCGCAGC--CTTCTCTCTCTCTCTCTCTCTCTCTCTCT 555
Oy 10934 GCTCCCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11013
Db 554 GCTCCCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 495
Oy 11014 TCCTCTGAGACCGCTCTCAAGAGCTCCACATCTCTCTCTCTCTCTCTCTCTCTCTCT 11073
Db 494 TCCTCTGAGACCGCTCTCAAGAGCTCCACATCTCTCTCTCTCTCTCTCTCTCTCTCT 435
Oy 11074 TGCAGTACTCCAGATGGGGCCCAAGAGGAGTAAAGAGGAGCAATACACTTCTCTCA 11133
Db 434 CACAGTACTCCAGATGGGGCCCTCACAAGAGGAGTAAAGAGGAGCAATACACTTCTCTCT 375
Oy 11134 CCCTGCTGAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11193
Db 374 CCCTGCTGAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 315

```

```

RESULT 15
LOCUS GGVITCR1 3297 bp DNA linear VRT 29-NOV-1987
DEFINITION Chicken vitellogenin gene 3' flanking region.
ACCESSION Y00324
VERSION Y00324.1 GI:63883
KEYWORDS CRI repetitive sequence; vitellogenin.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 3297)
AUTHORS van het Schip,F., Samalio,J., Meijlink,F., Gruber,M. and AB,G.
TITLE A new repetitive element of the CRI family downstream of the
chicken vitellogenin gene
JOURNALS Nucleic Acids Res. 15 (10), 4193-4202 (1987)
MEDLINE 87230999
PUBMED 3035488
FEATURES
source 1..3297
location/Qualifiers
/organism="Gallus gallus"
/db_xref="taxon:9031"
/cldone="24"
/clone_lib="lambda Charon 4A"
2221..2693
/note="repetitive element of chicken CRI family"
2413..2693
/note="resembling avian retroviral long terminal repeat"

```

```

BASE COUNT 738 a 809 c 981 g 769 t
ORIGIN
Query Match 3.1%; Score 375.2; DB 5; Length 3297;
Best Local Similarity 77.9%; Pred. No. 1.3e-77;
Matches 514; Conservative 0; Mismatches 118; Indels 28; Gaps 4;
Oy 10554 AAAACCTACTTATATCAATCATATCATATGATGAGGCTGGGTGGAGGAGGCCCAAG 10613
Db 2691 AAAAATAATGACAGAAATGACAGAAATGACAGAAATGAGCTGGGTGGAGGAGGACCTCGAG 2632
Oy 10614 GATCATGAAAGATCCAAACACCCCGCCACAGAGGGGCCAACCAACCTCCAGATCTGCTACT 10673
Db 2631 GATTCAGATCTCTCCAAACCCCTGCG--TGGGCAAGGGCCACCAACATCCCAT--TTACT 2577
Oy 10674 AGACGAGGAGCCGAGGAGGCTCCATCCAACTGGCCATGAAACCTCCAGAGGATGAGCAT 10733
Db 2576 AGACAGAGTGGCCAGGGGCCCATCCAACTGGCTTGAAACCTCCAGAGGATGAGCAT 2517
Oy 10734 CCACAACTCTCTGGGCAAGCTGTGCGACAGACCTCACCACCTCTCTGTGAAGAACTTTT 10793
Db 2516 CCACAACTCTCTGGGCAAGCTGTGCGACAGACCTCACCACCTCTCTGTGAAGAACTTTT 2457
Oy 10794 CCCTGACATCCAAATGAGCTCCCTGATGAGGTTAGATGATGATGATGATGATGATGAT 10853
Db 2456 CCCTGACATCCAAATGAGCTCCCTGATGAGGTTAGATGATGATGATGATGATGATGAT 2397
Oy 10854 CACTGTCTACTCTGTGTAAGAAAGTGTATCTCTCTCTT-----T 10893
Db 2396 TGTATCTACCTCTTCAAGAGTTTACTCCCTCTCTCTGTTATAGGCTCCCTTCAGGAGT 2337
Oy 10894 GGAAGGTTCAGATGAGTCTCTCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10933
Db 2336 GAAAGGCTGCAATGGGATCAGCGCGCAGC--CTTCTCTCTCTCTCTCTCTCTCTCTCT 2280
Oy 10934 GCTCCCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11013
Db 2279 GCTCCCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Oy 11014 TCCTCTGAGACCGCTCTCAAGAGCTCCACATCTCTCTCTCTCTCTCTCTCTCTCTCT 11073
Db 2219 TCCTCTGAGACCGCTCTCAAGAGCTCCACATCTCTCTCTCTCTCTCTCTCTCTCTCT 2160
Oy 11074 TGCAGTACTCCAGATGGGGCCCAAGAGGAGTAAAGAGGAGCAATACACTTCTCTCA 11133
Db 2159 CACAGTACTCCAGATGGGGCCCTCACAAGAGGAGTAAAGAGGAGCAATACACTTCTCTCT 2100
Oy 11134 CCCTGCTGAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11193
Db 2099 CCCTGCTGAGGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040

```

Search completed: March 27, 2003, 10:49:12
Job time : 22462 secs

SEARCH REQUEST FORM

Requestor's

Name: _____

Serial

Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

BEST AVAILABLE COPY

BEST AVAILABLE COPY

STAFF USE ONLY

Date completed: 03-28-03

Searcher: Beverly 64994

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 23

Number of Searches: _____

Number of Databases: 1

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG Suite

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

✓ Other CGN

THIS PAGE BLANK (USPTO)